THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jeffery William Bond

TITLE OF THESIS: A Descriptive Analysis Using the Coaching Behavior Observation System.

DEGREE FOR WHICH THESIS WAS PRESENTED: Master of Arts

YEAR THE DEGREE GRANTED: Fall, 1979

THE UNIVERSITY OF ALBERTA

A DESCRIPTIVE ANALYSIS USING THE COACHING BEHAVIOR OBSERVATIONAL SYSTEM.

by

© JEFFERY BOND

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

DEPARTMENT OF PHYSICAL EDUCATION

EDMONTON, ALBERTA

FALL, 1979

75F-25

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "A Descriptive Analysis Using the Coaching Behavior Observational System", submitted by Jeffery William Bond in partial fulfillment of the requirements for the degree of Master of Arts.

DEDICATION

To E. T., a Master Coach

ABSTRACT

The purposes of the study were to examine the validity of the categories of the Coaching Behavior Observational System (CBOS), to suggest guidelines for measuring observer reliability during training and behavior coding, to begin the development of a data base of video-tape recorded practice session coaching behaviors, and to investigate the usefulness of the CBOS descriptive analyses.

The coaching behaviors of two coaches were videotaped, coded, and analyzed to provide descriptive information concerning category frequencies, behavior clusters, coaching style indices, and sequential aspects of the observed behavior.

The findings support the utility of the observational method, and in particular the CBOS, for research on the coaching process.

It was proposed that intra-observer accuracy and stability, and inter-observer agreement measures be included in the observer training and coding procedures of the CBOS. The results of the various reliability measures reveal that the CBOS can be used in an accurate and consistent manner to provide useful descriptive information.

Modifications were proposed to enhance the validity of the CBOS category definitions in terms of the mutual exclus-iveness of the categories and the range of behaviors

accounted for by them. These included a number of additions to the category definitions, and suggestions for the further subdivision of the Instruct category.

ACKNOWLEDGEMENTS

I am greatly indebted to my thesis chairman, Murray Smith, for his untiring guidance and support, and to the other members of my committee, Len Wankel and Clare Drake, for their assistance with the preparation of this thesis. I have been fortunate to have had the benefit of the wealth of experience of these excellent teachers and coaches.

I am also grateful to two fine coaches for their contribution to the study, and for their perseverance with my presence during their practice sessions.

To my fellow graduate students and friends, and especially to Don Horwood and Barry Gibson, I offer my sincere thanks for continued support, advice, and encouragement.

Finally, the unending support of my wife, Patricia, our children, Michaele and Shaun, and our families, made my work possible.

"When a man does what needs to be done, he
does not know the meaning of time."

Chief Dan George.

TABLE OF CONTENTS

LIST OF TABLES

TABLE		PAGE

LIST OF FIGURES

CHAPTER I

STATEMENT OF THE PROBLEM

Introduction

Sporting involvement in Canada, as in other western countries, is truly on a large scale. For example, the Canadian Amateur Hockey Association registration figures for minor hockey in the 1977-78 season show 245,120 players in the peewee (12 years old on January 1) and categories below, 108,550 at the bantam (13 and 14 year olds) level, 84,400 at the midget (15 and 16 year olds) level, 33,300 at the juvenile (17 and 18 year olds) level, and 11,800 at the junior level (Millson and Christie, 1979). These figures do not include players in recreational leagues that do not affiliate with the governing body, nor do they include players in high schools, colleges and universities. It is important to note that the lion's share of organized hockey for all ages has been handled by community clubs outside of the school systems. Very few Canadian junior or senior high schools maintain quality programs in competitive hockey. Basketball is the major school sport over the winter months. For teams affiliated with the C.A.H.A. the writer estimates that the number of minor hockey coaches probably exceeds 20,000.

At the provincial level, the Executive Director of the Alberta Schools Athletic Association estimates that approximately 24,000 children and 1,200 coaches are

involved in a variety of sports at the interschool level in the 265 affiliated high schools (grades 10, 11 and 12). These figures do not include students and coaches involved at the intermural level (Wood, 1979).

Although it is difficult to obtain accurate figures on the total number of coaches across all sports, the 1979 registration records of the National Coaching Certification Program in Alberta show that approximately 5,000 coaches have enrolled in at least one level of the theory or technical components of the certification program. This represents approximately 10 percent of the estimated total number of volunteer community coaches in Alberta (Fodor, 1979).

Even rough estimates of participant and coach involvement across Canada based on the Alberta estimates indicates a need for continuing inquiry as to the purposes, content, and effectiveness of youth sports and coach education programs.

It is evident that there is some conflict over the legitimate purpose of youth sports. There are those who claim that the outcome of youth sport programs must be designed to meet the needs of the participants (Smith, 1973,1974a, 1975a; Orlick, 1974; Orlick and Botterill, 1975), and there are those who accept that the outcomes are designed to meet the needs of the increasing number of adults involved in the coaching, organization and administration of youth sports. Sport must inevitably

do both; but it must be the needs of the participants that are given priority. It must also serve the needs of adult organizers and coaches, but those considerations must be made to compliment and not supercede the needs of participants (Smith, 1974a; Orlick and Botterill, 1975; Devereux, 1976).

There is little doubt that the quality of the experience offered to the young sports participant is directly related to the total sporting environment with which he or she is associated. The total sporting environment may be seen to consist of coaches, administrators, supporters, officials, other players, and the physical characteristics such as equipment, playing surfaces, and so on. In reference to the quality of the sporting experience, Orlick (1974) states:

> Sports are not inherently or necessarily good, as was once supposed, nor are they inherently bad; rather they have the capacity to be either, depending upon the models and reinforcement contingencies which are operational for the particular individual within and around the specific sporting environment (p. 2).

One important part of the sporting environment is of course the coach, however, until the last few years very little has been offered in the way of coach education courses. As Smith (1973, p. 1) has pointed out "The overwhelming majority or organization and coaching at *all* levels in sports or physical activity in Canada is carried out by lay volunteers". If the figures presented above are an accurate indication the position regarding

approaches to training coaches in Canada, as in many other countries, has not changed a great deal since 1973.

In the majority of cases the local community coach has probably, but not necessarily, had some years of competitive experience as a player and upon retiring or reaching a stage of life where he or she feels that sufficient time is available, he or she seeks a coaching opportunity in the community. The average coach brings with him or her the experiences gained as a player and the memories he or she has of the training procedures, skill practices and strategies that were involved. In addition to this experience, the coach usually possesses intentions regarding the improvement of his players' skills and game performance, either for the players' benefit, or his or her own. However, some previous playing experience and "good" intentions may not be sufficient in terms of providing satisfying experiences for the participants. What other alternatives exist for the improvement of coaching skills?

The National Coaching Certification Program in Canada currently offers courses at three levels in the theory of coaching, covering general aspects of coaching common to all sports. At each of the three levels, the participant is introduced in successively more detail to: the role of the coach, sports psychology, motor learning and motivation, growth and development, biomechanics, exercise physiology, sport medicine and training methods. The level one theory course requires the participant to attend lectures and

discussion for 16 hours.

A second component of the Certification Program is concerned with the technical aspects of coaching which are specific to the various sports. These technical courses are prepared by task-force groups designated by the specific sport governing bodies in accordance with broad guidelines set down by the coordinating body of the national program. The technical aspects covered by these courses include the various skills, drills and strategies appropriate to the specific sports.

The final requirement for a coach to obtain certification at a particular level is a specified amount of practical experience ranging from one year at level one to five years for level three.

It is important to note that, at present, there is no intention of eliminating amateur coaches who do not gain certification through this program. However, the level at which certification has been attained is becoming an increasingly significant factor in selecting coaches for the more attractive coaching assignments.

Although this approach to coach education is to be commended for its thorough planning and coordinated administration, it is not without its critics. Issues that remain to be solved include the amount of time that a volunteer community coach should, and is prepared to set aside for training programs of this type, and the inclusion and content of coaching methodology in theory

courses. There would appear to be a case for arguing that a considerable amount of time in coach education courses is spent on telling coaches "what", and insufficient time is spent on telling coaches "how" to coach. Usher (1977) states that in some cases any special knowledge of pedagogy as a basic science of coaching is felt to be unnecessary. It is often argued instead that a "good" coach is simply one who knows his subject matter and is interested in it.

A second alternative which exists in Canada for those who wish to improve their coaching skills is to be found in the various coaching routes offered by colleges and universities, usually as part of the undergraduate physical education program. At the University of Alberta for example, the coaching route requires students to undertake a number of compulsory courses on the theory and practice of coaching in association with individual and team sport activity, as well as coaching practicum courses. The coaching work includes both theory and technical components of the National Coaching Certification Program.

Despite these two alternatives two facts remain; the first is that the vast majority of coaches are untrained in terms of some type of formal coach education course, the second is that we have very little information available concerning the coaching process. Smith, Smoll, Hunt, Curtis and Coppel (1978, p. 3) state: "Although emphasis is placed on the importance of 'high quality supervision',

research pertaining to coaching behaviors and their influence on children has been virtually non-existent".

We need to know about the nature of the coaching role; what a coach does, how he influences skill learning, how he influences an athlete's motivation and performance, what aspects of a coach's behavior exert influence on children's attitudes toward the coach, toward sport participation, and toward themselves. If one accepts the view that youth sports exist for the participants and that sport is an additional means (other than home and school) of assisting self-growth and self-discovery for the participant, then it is important that researchers examine the basic processes of coaching and provide input relevant to coaching methodology in coach education courses (Smith, Smoll and Curtis, 1979).

A good deal of the research on coaching has done little to provide us with information on the issues mentioned above. The approaches adopted have been in Ritzer's (1975) terms "multiparadigmatic". Ritzer bases his discussion of multiple paradigms on Kuhn's (1962) proposal that science makes its great advances during periods of "revolution", rather than during the more lengthy periods of "normal" activity. (It is not intended here to enter into a discussion of the characteristics of a "science", however, for a succinct discussion of the relationship between physical education, the sport sciences and scientific method, see Smith, 1975b).

Kuhn's model of scientific change is as follows: Paradigm 1 $\longrightarrow$ Normal Science $\longrightarrow$ Anomalies $\longrightarrow$ Crisis $\longrightarrow$ Revolution $\longrightarrow$ Paradigm 2, where a paradigm is seen as a "fundamental image of the subject matter within a science" (Ritzer, 1975, p. 7). According to Ritzer, multiple paradigms result in it being almost impossible for researchers to get down to the business of "normal" science and thus inevitably to "revolution" and advance.

The failure of the multiparadigmatic research on coaching to provide us with useful information about coaching behavior rests partly with researchers who are constantly defending their own position and assumptions, and partly with those researchers who are unable, or not concerned with developing what Smith (1975b, p. 230) has described as: "an accurate perspective of the modern world along with the integrative abilities required to relate sophisticated technical knowledge to its real world context".

The trait or sport personology approach has attempted to provide us with a profile of the average or "good coach" based on certain personality characteristics or traits (Ogilvie and Tutko, 1968; Hendry, 1969). Given the controversy which is generated by the particular personality theory adhered to by different researchers, and the differences in research methodology, it is apparent that the trait approach is somewhat limited in terms of the specificity of description and the accuracy of prediction

(Blumer, 1967; Kroll, 1970; Martens, 1975; Fisher, 1977; Fisher, Horsfall and Morris, 1977).

The situationist approach (Rushall, 1978) has pointed to the importance of environmental or situational factors in determining coaching behavior. The normative model of decision styles in coaching presented by Chelladuria and Haggerty (1978) is an attempt to relate decision styles to certain environmental constraints such as restrictive time pressure, a requirement for quality of performance, problem complexity, and so on. However, just as it is illogical to conceive of behavioral variation as being totally attributable to traits, it is no less illogical to over emphasize situational characteristics.

While most advocates of situation specificity do not completely discount individual personality differences, and most supporters of trait personology do not dismiss environmental factors entirely, some sources have indicated the mutual involvement of traits, situations, and the interactions between these, in the determination of behavior (Endler and Hunt, 1966; Endler and Magnussen, 1976). This interactionist position is also not without its critics (Cartwright, 1975; Golding, 1975; Morgan, 1978).

The point to be made here is that we are left without a great deal of useful information about the coaching process. The beginning coach is confronted with a good deal of contradiction and controversy and a sometimes bewildering array of research approaches and data. There

has been very little research on the characteristics of coaching behavior at various sporting levels and across a variety of sports.

Teaching and coaching have a good deal in common. This point is discussed in more detail in the following chapter; but a statement by Simon and Boyer regarding teaching techniques is equally applicable to sports coaching, and is relevant to the issue under discussion here:

> A great deal has been written lately about teaching techniques designed to help teachers engage in different types of classroom processes. Much of the literature about these classroom processes does not specify the teaching strategy necessary in order to be able to bring about desired outcomes. Often, this is because the tools for discussing the necessary teacher behaviors are not available, or not known about, by the curriculum authors. Thus, the desired outcome for students is discussed, but the way to get there is not (1974, p. 5).

The more recent emphasis on interactionism in sport psychology research has pointed to the need for the observation or measurement of behavior in the natural setting. A growing number of researchers (Smith, 1978; Smith, Smoll, Hunt, Curtis and Coppel, 1978) have moved toward observational analysis as a research tool to improve our understanding of what coaches actually do. Smith states:

> There are many possible approaches to such a study, but any approach would be enhanced by a tool that would enable one to identify the most important behaviors, categorize these behaviors in a functional way, and use this

> information as the basis of an analysis that would provide insight into the nature and effects of communication (1978, p. 38).

The observational method is described by Weick (1968) as "...planned, methodical watching that is delimited by constraints to improve accuracy" (p. 358). The relative recency of the use of observational analysis in the study of coaching behavior requires firstly that much of the initial thrust of the research effort be directed at appropriate instrument development, and secondly that the substantial literature already available from research on classroom teaching be carefully studied and utilized as a basis for that development.

Purposes of the Study

The purposes of this investigation are to begin to establish a data base of videotape recorded interactions between coaches and athletes, and to categorize and analyze coaching behaviors using the Coaching Behavior Observational System (CBOS) developed by Smith (1978).

More specifically, the objectives of the study are:

1. To examine the CBOS category definitions and ground rules with regard to the content validity of the categories, the explicitness of the ground rules, and the associated reliability of observation. Since actual instruction is regarded as one of the

central functions of coaching, the important components of behavior included in the rather global category of Instruct as presently defined in the CBOS, are to be identified and the category further subdivided to accommodate these various sub-classes of behavior.

2. To study the usefulness of grouping and analyzing the data in terms of the time related measures, behavior clusters, coaching style indices, and behavior sequences suggested by Smith (1978), and to modify and develop alternatives to these where this might be indicated.

3. To formalize a set of training instructions appropriate for an investigation of this type, and to provide guidelines for reliability measurements important in establishing the accuracy and consistency of observations.

Nature of the Study

Limitations

This study represents a counterpart to the early stage of a sequence in research on teaching suggested by Anderson (1971) which begins with the development of systems for describing events in teaching [coaching]

settings, and ends with the utilization of information gained through research to enrich teacher [coach] education programs. It is a descriptive rather than evaluative or prescriptive approach which focuses upon coaching behavior; it is not concerned with the interactions between athlete and coach, or with the effects of coaching behavior on player attitudes. It is understood that athlete behavior is important and that interaction between coach and athletes is the essence of coaching. These matters will have to be dealt with in future studies after appropriate development of the means for accurately describing coaching behavior.

It is important to recognize that the presence of an observer and videotaping equipment may have some influence on coaching behavior directly, or indirectly by influencing player behavior and the coach's reactions to it. Johnson and Bolstad (1973) state that the presence of these reactive effects depends upon many factors including the setting, the length of observation, the constraints placed on subjects by the conditions of observation, the conspicuousness of the observer, individual differences between subjects and so on. They suggest that a thorough rationale for being observed may assist in reducing reactive effects. While an attempt has been made in this study to reduce reactive effects by discussing the rationale with the coaches involved, by arranging familiarization visits to a number of practices

prior to the commencement of data collection, by reducing observer conspicuousness, and by the sampling procedures employed, the possibility of observer influence remains as it does in all studies where a known observer and equipment are present.

A second type of reactivity is related to observer reactions to reliability checks. Johnson and Bolstad (1973) have mentioned such factors as heightened motivation within the observer for accuracy, heightened vigilance for critical behaviors, and heightened perception by the observer for the coding idiosyncracies of the calibrator. In addition, Kazdin (1977) proposes that during reliability checks observers normally code with reduced complexity which usually leads to higher agreement scores. Kazdin suggests various strategies for overcoming or reducing observer reactivity, including the use of different assessors, leading observers to believe that all of their observations are to be monitored, and making reliability checks as unobtrusive or covert as possible. Kazdin refers to situations where there are a number of observers and possibly a number of different reliability assessors. These situations did not exist in this investigation, and although a number of steps were taken to accurately measure and maintain related aspects of reliability, the possibility of observer reactivity cannot be entirely discounted.

Another phenomenon which may contribute to a

reduction in the consistency and accuracy of observation has been called "observer drift" (Kazdin, 1977). Training procedures are designed to ensure that observers adhere to the category definitions and ground rules and record or code behavior at a consistent level of accuracy. Recent evidence cited by Kazdin supports the concept of observer drift which is associated with the length of the investigation and the explicitness of the category definitions and ground rules. Drift may be reduced by continuous training and discussion, accuracy feedback, and the random analysis of videotaped data. Several of these safeguards were adopted in this study, although it must be accepted that accuracy and consistency cannot be perfect or remain static. The alternative is to accept certain variations in accuracy and consistency within a commonly accepted bandwidth.

Delimitations

This investigation is delimited to a preliminary description of the coaching behavior observed and videotape recorded during practice sessions of two university level coaches involved in the coaching of university athletes during the winter term, 1979, at the University of Alberta; to the use of the CBOS; and to an analysis of these coaching behaviors. The CBOS is designed for the analysis of coaching behaviors during practice or workout sessions, rather than during games or contests, or lectures

and meetings (Smith, 1978, p. 1). As such this study is concerned with a limited, but nonetheless important range of coaching behavior.

While the coaching level employed in this study has been delimited, so too has the sporting activity. The selection of university level basketball and volleyball was partly a function of the availability of activities and adequate facilities for recording coaching behavior.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

In the previous chapter some comment was made about similarities between teaching and coaching, and the apparent lack of information concerning the coaching process or methodology. It was suggested that the findings from classroom observational research be used as a basis for the development of instruments appropriate for the analysis of coaching behavior. Because some of the following discussion involves information derived from the considerable body of classroom research, and the intention is to support the adoption of the proposals and suggestions from that research in the investigation of coaching behavior, it is appropriate at this point to take up further discussion of the teacher-coach classification.

One of the essential functions of interaction in instructional groups is that of communication; that is, the communication of cognitive (technical), affective (emotional), psychomotor (non-verbal), procedural and other types of information in a reciprocal fashion between leader(s) and member(s) or member(s) and member(s) of the group. The main purpose of this interaction is to arrange the learning environment in a way that is conducive to optimum levels of motivation, learning and performance so that self-growth and self-discovery or personal development by the learner will be enhanced. This characteristic

of interaction is a fundamental feature of both the classroom and sports field instructional setting and serves as a basis for the development of category systems for observational instruments.

Bany and Johnson (1964, pp. 33-38) discuss a number of features which are characteristic of groups in general, and applicable to classrooms and sports instruction settings. Interaction in groups refers to the ways in which positive, neutral, and negative influences are exerted between members, sub-groups within the group, and between leaders and groups. Groups generally possess some type of system of social or other stratification, and can be characterized in terms of some degree of cohesiveness which develops over time, partly as a response to shared motives and goals.

Smith (1978) suggests that coaching is a special form of teaching, although he does point out some differences:

> Briefly, coaching differs from teaching in that it usually involves fewer individuals; goals are more likely to be shared by participant and coach; the relationship is more intense and extended; and, performance criteria are usually more demanding and objective (p. 44).

Indeed, it is likely that classroom teachers would agree that their effectiveness would be increased if they could work for more extended periods of time, with smaller groups of learners, among whom there was more agreement as to goals and a strong committment to higher standards. In this sense, coaching might well be described as a

form of teaching in which some highly desirable conditions prevail.

The similarities between teaching and coaching discussed by Smith involve a primary concern with the learner's acquisition of knowledge, skills, problem-solving ability, and social-emotional maturity. It is this common ground that justifies the need for a close look at classroom observational research by those concerned with the study of coaching behavior, in terms of the relevance of its findings and the appropriateness of applying classroom instruments to the observation of coaching behavior. These aspects are discussed in the following section of this chapter.

Observational Analysis and the Observational Method

Simon and Boyer (1974, p. 3) speak in terms of a "meta-language" for describing communication. A meta-language is seen by them as important in a profession where communication must be accurately transmitted from one practitioner to another. They suggest three necessary requirements for such a language. First, it must be descriptive rather than evaluative; second, it must deal with what can be categorized and measured; and third, it must deal with small bits of action rather than global concepts (p. 4).

A useful discussion of the observational method is provided by Weick (1968). He states:

> An observational method is defined as the selection, provocation, recording, and encoding of that set of behaviors and settings concerning organisms 'in situ' which is consistent with empirical aims (p. 360).

Each of the key words is examined in further detail by Weick. For example, "selection" refers to the choices that observers make regarding the inevitable editing or focusing that takes place before, during, and after observation. The word "provocation" suggests that observers can, and should, make subtle changes to natural settings to increase clarity, without destroying the essential nature of the setting. The term "in situ" indicates that observations are typically made in situations which are normal or familiar to the subjects. Finally, "empirical aims" suggests that observational analyses can be directed toward hypothesis testing or formulation, as well as description.

Flanders (1965, 1970) further emphasizes the systematic nature of observational analysis, its attempts to minimize bias on the part of the observer, and its scrutiny of the process of instruction by taking into account each small bit of interaction. With regard to this latter feature, Martin (1976) states:

> By breaking up the continous stream of teacher-student behaviors into readily observable units, such instruments permit the production of permanent 'short hand' records of classroom interactions. These records can be analyzed by teachers and research workers in a variety of ways and for a variety of purposes (p. 5).

The *raison d'être* for observational instruments is to aid direct observation. If one really wants to know about a particular phenomenon which exists in nature then one can understand and learn about it by direct observation. However, this direct observation requires direction and control if it is to produce valid and reliable measures. Anderson (1972) supports this view:

> These techniques are needed to delimit the observers task, as it would be impossible to observe everything that takes place in an instructional setting. These techniques are also needed to obtain a valid and reliable conception of what is being observed (p. 4).

There are various types of observational instruments which have been distinguished in terms of three main elements by Rosenshine and Furst (1973). The first of these elements concerns differences in recording procedure. *Category instruments* record an event each time it occurs within specified time intervals. For example, for each five second interval an observer encodes the various behaviors according to the category definitions of the instrument. Martin (1976) states:

> These instruments typically record classroom behaviors in the form of tallies, checks, or other marks which code them into pre-defined categories and yield information about which behaviors occurred and how often they occurred during the period of observation (p. 5).

Sign systems require the observer to record a behavior that occurs, regardless of its frequency of occurrence.

Rating systems involve the estimation of the frequency of specified events, usually at the end of an observation session.

The second main element of distinction discussed by Rosenshine and Furst concerns differences in items. Traditionally, rating systems were characterized by the inclusion of high inference items which may be conceived of as broad or global items requiring much inference by the observer, while category systems and sign systems typically contained low inference or specific items. An example of a high inference item is a rating of overall teacher effectiveness, while that of a low inference item is the coding of behavior associated with the teacher giving directions. Rosenshine and Furst note, however (p. 133), that each of the three types of instruments may now be found to contain high, moderate, and low inference items.

The third element by which distinctions can be made between instrument types is related to differences in the format of coding, from simple uni-dimensional to multi-dimensional coding.

Biddle (1967) also includes the unit of analysis as a further way of describing various instruments. He states that an instrument may use an arbitrary unit of time such as a five second or five minute interval, selected naturally occurring events such as observing a teacher only during periods when he is lecturing, analytic units

where for example, a unit may be one or more verbal exchanges that comprise a completed verbal dramatization, or phenomenal units which are designated as natural breaks in the stream of classroom processes.

While the above basis for the distinction of various observational instruments are certainly not exhaustive, they do indicate the many and varied approaches which have been taken. According to Rosenshine and Furst (1973) there are well in excess of 120 classroom observational instruments. Their main source for this estimation is the anthology by Simon and Boyer (1967, 1970) which contains 92 systems.

Some Essential Features of Category Systems

Martin (1976, p. 5) has described some essential features which are of primary concern for category instruments of the type used in this study. These include a set of operationally defined categories of behavior, a set of rules and priorities for observation and coding, a standarized recording form, and a series of instructions for organizing and analyzing the observational data. "A sound category observation instrument must be objective, relevant, parsimonious, efficient, reliable, and valid" (p. 6). In his discussion of the qualities of a "sound" instrument, Martin notes that to be objective, an instrument must contain categories which are operationally defined in such a way as to insure that the behaviors

recorded are directly observable rather than hypothetical constructs. The adequacy of this definition is of course related to the reliability of observation. Relevance is closely linked with the purpose for which the instrument is devised, and is associated with instrument validity. Parsimony refers to the complexity of an instrument in terms of the number of categories, and the coding procedures employed. Efficiency is related to the ease with which one can distinguish between observed behaviors according to the various category definitions and ground rules. Efficiency is thus important in terms of observer training and the generality of users of the system. The final essential characteristic discussed by Martin, that is, reliability and validity, deserves further expansion.

In classical test theory, the concept of reliability involves the consistency with which a test measures a given attribute or yields a consistent score on a given dimension. Reliability, as usually employed in observational research, concerns two separate features broadly described by Kazdin (1977) as observer accuracy and observer agreement.

Accuracy refers to the extent to which the observations scored by an observer match those of a predetermined standard for the same data. The standard used for the accuracy test may be determined by several observers who reach consensus, or by using specially

constructed observational material.

Observer agreement may be further divided to include inter-observer and intra-observer agreement. *Inter-observer agreement* between observers who independently code the same behavior of the same subject indicates that the observations reflect the subject's performance relatively accurately, and that both observers agree on the interpretation of the category definitions and ground rules at that time. *Intra-observer agreement* is calculated on two or more repeated observations by a single observer of the same behavior of a single subject, with the repeated observations separated by an interval of time. This of course necessitates the use of videotaped records of the subject's behavior which can be coded as many times as required to measure intra-observer stability. An acceptable measure of intra-observer stability provides an estimate, as to the adequacy of the consistency with which the observer is coding the same behavior at different times, and represents a measure of the observer's stability of interpretation of the category definitions and ground rules.

Hollenbeck (1978) notes that inter-observer agreement does not by itself assess observer accuracy unless a previously established standard is used, and does not assess stability unless it is measured over repeated trials. Therefore, it is possible to obtain high inter-observer agreement with close to zero reliability

in terms of accuracy and stability. As Hollenbeck points out: "...all agreeing observers could simply be applying the same incorrect behavioral definitions" (p. 4).

Poor reliability, resulting when two measures of the same subject differ considerably, can occur because:

> ...the behaviors are unstable, because the observers are unable to agree on what occurs, because the different items which enter into the measurement lack consistency, or for some other reason (Medley and Mitzel, 1963, p. 291).

Some of the "other reasons" have been previously discussed in the first chapter. These include observer bias, observer drift, and reactivity on the part of subjects or observers.

Inadequate reliability measures may result because the investigator has not properly distinguished between the three types of reliability. Most authors who have reviewed observational measurement issues have dealt with the concept of reliability, but there is often no consensus in the terminology, and no agreement as to the statistical approach which should be adopted.

Overall percent agreement is the more commonly used measure (Johnson and Bolstad, 1973; Hawkins and Dotson, 1975; Kazdin, 1977; Hollenbeck, 1978), with an agreement score of 80 to 85 percent considered adequate. The traditional method of calculating percent agreement is to divide the number of agreements by the number of agreements plus the number of disagreements and multiply

this by 100. An agreement refers to any interval in which both observers recorded that a behavior occurred or in which both observers did not record the behavior; a disagreement refers to an interval in which only one observer recorded a behavior.

Hawkins and Dotson (1975) suggest at least three potential sources of error in obtaining accurate and objective data. The first is that the definition of the behavior category may be vague or involve high observer inference, or that the behavior may be difficult to detect because of its complexity or subtlety. The second is that the observer may be poorly trained, lack sufficient motivation, or be otherwise incompetent. The third is concerned with observer bias and related issues. Hawkins and Dotson have succeeded in demonstrating the inadequacy of overall percent agreement in terms of detecting these three sources of error, in a series of simple, but ingenious studies. They propose two alternatives; the first involves a complex mathematical solution such as a set of tables based on the probability of an obtained difference between two observers' data, given the particular frequency of the behavior; the second involves alternative formulae for calculating various aspects of observer agreement and disagreement.

Hollenbeck (1978) briefly reviews various statistical approaches to reliability including several variance analyses and nonparametric procedures, and although he

finds strengths and weaknesses in the majority of them, he proposes the use of an agreement coefficient suitable for nominal data which makes a correction for chance agreement. He refers to Scott's (1955) coefficient and Cohen's (1960) Kappa as two procedures which satisfy these requirements.

Although Hollenbeck suggests the use of Cohen's Kappa, the Scott coefficient has been chosen for this study because it has been widely used in observational research and is often recommended for its ease of calculation. Flanders (1967) states:

> Scott's method is unaffected by low frequencies, can be adapted to percent figures, can be estimated more rapidly in the field, and is more sensitive at higher levels of reliability (p. 161).

Validity is commonly subdivided into different types. Quereshi (1974) and Cronbach (1970) for example discuss the features of face, content, criterion, and construct and validity. Each of these has relevance for observational research, but as is the case with reliability issues, very little is reported in the literature (Hollenbeck, 1978).

Face validity refers to the outward characteristics of the observations, measurements or interpretations which, according to Cronbach (1955), should have relevance to the layman and should appear to be relevant to the particular conditions under which the observations are made. Face validity on its own is often insufficient or inadequate

without support from other types of validity.

Content validity is ordinarily established deductively by defining the universe of items and sampling systematically within this universe to arrive at category items. In observational research this is often accomplished through content analysis.

Criterion validity, which Cronbach (1955, 1970) further subdivides into predictive and concurrent validity, involves a correlation between measures obtained by an instrument and some form of independent criterion. In observational research the measures of behavior may be correlated with independent measures of the same behavior obtained through questionnaires or interviews; but this is difficult in practice because of the problems associated with getting an independent measure which uses the same dimensions of behavior as are sampled by the instrument.

Construct validity concerns whether or not a particular construct applies to a measuring instrument or interpretation. In this case it is necessary to derive hypotheses about the behavior of interest, or about the instrument, from theory related to the construct, and to verify them experimentally (Cronbach, 1970). This type of validity is seldom researched in observational studies.

Medley and Mitzel (1963) have stated what they believe to be the validity which is applicable in observational research as: "...the extent that differences in scores yielded by it reflect actual differences in

impressions made on different observers" (p. 291).

This is extended by Anderson (1972) who states:

> For a category system to be valid it was considered that a representative sample of the behaviors must be observed, an accurate record of the observed behaviors must be obtained, and the records must be scored so as to reflect faithfully differences in behavior. From the above it can be seen that the reliability of the observer using the category system has an effect on the validity of the system (pp. 16-17).

The Uses of Observational Research

Observational systems have been designed for a variety of both general and specific uses. Even within the field of physical education observational instruments have been used for a variety of specific purposes. In general terms, observational analyses have as their main purposes:

1. To describe current practice, be it in coaching or teaching. For example, concern might lie in describing variability in behavior over time for the same individual, between individuals and situations, or in describing characteristic behaviors of superior coaches or teachers (Rosenshine and Furst, 1973; Smith, 1978).
2. To train teachers or coaches to learn to identify and describe fundamental behaviors and the ways in which these are combined or patterned (Smith, 1978; Smith et. al., 1978).

3. To monitor instructional programs (Smith, 1974; Smith et. al., 1978).

4. To investigate relationships between classroom or coaching activities and measures of participant growth (Rosenshine and Furst, 1973).

The Development of Category Systems

Information and suggestions regarding the development of category observation instruments has been presented by Biddle (1967), Fishman and Anderson (1971), Rosenshine and Furst (1973), and Smith (1978) among others. A concise and adequate treatment is provided by Martin (1976) who proposes a five step sequence for instrument development. A summary of this sequence is presented here as an example of the developmental phases which the instrument used in this study (CBOS) passed through.

The first step in the sequence described by Martin is concerned with the development of a pool of descriptive phrases from representative samples of the interaction appropriate for the intended use of the instrument. This pool is then categorized by grouping the informal observations into a number of "generic blocks", described by Martin as:

> ...a finite number of generic behavioral blocks. These categories, when taken together, should be a parsimonious abridgement of the variance and substance of the initial 'raw' observations (p. 9).

A method which is commonly used to develop categories for observational systems is content analysis, referred to by Berelson (1952) as: "a research technique for the objective, systematic and quantitative description of the manifest content of communication" (p. 18). Bowers (1970) states that one of the chief advantages of content analysis is that it guards against selective perception and associated distortion of such analyses. He does point out, however, that selectivity is involved in choosing the sample and in establishing the category system used as a basis for the analysis; but that this approach often leads to efficiency of effort and additional information which may otherwise not be gained. Bowers suggests a sequence of interdependent steps to be followed during content analysis: development of hypotheses from hunches, common sense, extensions of elaborate theories or previous research, selection of the sample to be analyzed, selection of the categories to be used in the analysis, development of judgmental procedures for making decisions about the content of the sample, selection of a control or normative sample, reformulation of general hypotheses, selection of statistical tests, tabulation of data, and finally, application of the statistical tests. Simon and Boyer (1974) have grouped categories into seven classes: affective (primary focus is the emotional component of communication), cognitive (focus is the intellectual component of communication), procedure or routine (such as

getting ready to work, and administrative routine), physical environment (describes the physical space in which the observation takes place), psychomotor (focus is the description of non-verbal communication), sociological structure (provides information about those who are interacting), and activity (describes the type of activity).

An integral part of categorizing is the considerable use the instrument developer makes of the available findings from previous research in the area, in order to: "...capitalize on them and to avoid wasted effort in searching for answers already available" (Smith, 1978, p. 38). Martin (1976), in speaking about the instrument developer states:

> The greater his deductions coincide with the inductive work of others in the chosen area of investigation, the less likely a category constructor will waste his time 'reinventing the wheel' (p. 9).

A process which is compatible with the categorizing of this second stage in the sequence involves the selection of the unit of measurement so that the continuous behavior being observed can be coded in a meaningful and reliable manner by the observer. Martin calls this "unitizing". Precise ground rules must also be developed in the early stages to assist the observer in interpreting the category definitions.

The third step in the sequence is concerned with pilot testing the instrument in order to establish

acceptable reliability and validity. This is a progressive step which involves constant re-examination of the basic criteria which were noted earlier in this chapter (objectivity, relevance, parsimony, and efficiency). Acceptable reliability and validity must be generalized across observers and where applicable, across situations.

Formal categorizing and prioritizing are the objectives of the fourth step in the sequence. Martin describes prioritizing as:

> ...the process of developing rules on the basis of which an observer-coder can record 'more important' behaviors at the expense of 'less important' behaviors during complex or rapid interaction segments which make exhaustive recording impossible (pp. 11-12).

The final stage standardizes the procedures for recording and analyzing behavior, and for training observer-coders.

Although this sequence is presented here as an example of an adequate approach, it is not claimed that it is the best or only one which should be considered. It should also be noted that Martin points out that the stages of his sequence are not mutually independent; in practice there is considerable overlap and repetition.

Some Issues in Observational Research

Some of the problems associated with observational research and instrument development have already been touched upon in the previous pages. For example, the issues of reliability and validity, the wide number of

choices facing the potential observational instrument developer and user in terms of the large number and types of instruments available, the differences in items and recording procedures, and so on.

Little has been said thus far about the optimal number of categories which should be contained in an observational instrument. In discussing this issue Anderson (1972) notes that a high degree of reliability may be obtained by using small, easily observed, and easily recorded units; but this may render the resulting data as practically useless because the original continuity has been so severely disrupted. Broad category definitions may result in a high degree of validity, but low reliability because of the ambiguity of the definitions or the high observer inference required. However, it is generally stated (Martin, 1976) that a measure cannot be valid unless it is also reliable, and that reliability alone does not ensure validity. Martin (1976) calls for parsimony in selecting a relatively small number of categories. Smith (1978) states:

> It is necessary to strike a balance between a small number of categories that are so general as to provide no useful information, and a large number of categories that allow for many fine distinctions, but are cumbersome to use, very difficult to learn to apply, and pile up masses of data even from the examination or analysis of short episodes (p. 39).

Another important issue in observational analysis

concerns what should be observed. An observer may code behavior in terms of the intention, the overt structural characteristics, or the effects of the behavior. Biddle (1967) suggests that the selection of what to code depends to some extent on the objectives of the study. For example, if one is solely concerned with coach competence, then the effects of coaching behavior upon the participants may be more important. Flanders (1970) supports the coding of the objective characteristics of observed behavior. Coding on the basis of perceived intent is difficult, if not impossible, to do accurately and reliably. Hough (1967) and Medley and Mitzel (1963) argue that it is the effects of an action that result in precise and valid records, and that regardless of the intent, it is the effect on the recipient which is important. Smith (1978) also supports this view. Ground rule 2 of the CBOS states in part:

> Because the system is descriptive, not evaluative, the observer should not try to guess the intent of the coach, but rather to assign behaviors to categories according to how he or she perceives their <u>effect on the participants</u> (p. 6).

There are some specific issues related to the unique non-verbal and movement character of sports and physical education instructional environments, which face the researcher who intends using an available classroom observational instrument in movement settings.

Bookhout (1967) has stated that it is unsafe to assume that classroom observational instruments are

applicable to physical education, which occurs in: "...a broader, freer environment, has different subject matter and uses movement as its primary medium of learning" (p. 337). Anderson (1971) supports this view, arguing that an emphasis on verbal behavior, for example, as in Flanders (1965), limits the utility of classroom observational systems in physical education classes. He states: "...specialized systems will have to be developed, or existing systems modified to adequately describe events in physical education" (p. 6). Smith (1978) emphasizes the "natural unity" of language, thought and movement, and hence the need for both verbal and non-verbal categories. He also argues for the importance of a "monitor" category, as opposed to the "silence-confusion" category which is popular in instruments which analyze only verbal behavior. Monitoring by a coach or physical education teacher is an important part of the instructional process. Contemplation, structuring, mental rehearsal, and observation prior to correcting and so on are necessary features of activity instructional settings.

Observational Analysis in the Classroom

The review by Rosenshine and Furst (1973) provides an overview of the classroom observational research to that date. Of the 73 systems included from the 120 or so available at that time, 15 contain variables derived from "specified established theory"; 10 refer to theories or

research as the source of their variables; 24 are modifications of existing instruments, the most frequently cited being that of Flanders (1965); 24 are classified as being author-originated (pp. 138-146).

In terms of the purposes of particular instruments, Rosenshine and Furst found 36 designed to describe, analyze or observe classroom practice; 11 concerned with teacher training by providing a teacher with feedback on his behavior, a set of procedures with which he can categorize instructional activities, and behaviors and activities which he can model during instruction; 16 specifically designed to monitor instructional programs; and 7 for use in investigating relationships between classroom activities and measures of student growth. The balance of the 73 instruments included in the review were difficult to place in these categories.

By far the most cited instrument in terms of purposes and modifications is that of Flanders (1965). Flanders (1970) suggests that the inquiry behavior of teachers should include five steps:

1. Specifying the pupil behaviors described in class.
2. Identifying the patterns of teaching behavior considered likely to fit such pupil behavior.
3. Practising the teacher behavior patterns.
4. Designing a way to test the relationship

between pupil and teacher behavior.

5. Carrying out the plan in the classroom and testing the results.

An important point discussed by Rosenshine and Furst (1973) concerns the interpretation of the output of classroom observational research. There are two aspects to be considered, the first is related to the fact that although authors insist that their research is descriptive rather than prescriptive, as Rosenshine and Furst point out: "...the evidence of their reports has shown that certain variables become 'dos' and others become 'donts' fairly quickly "(p. 160). The second aspect of the interpretation issue concerns people other than the researcher, who read the research literature and judge descriptive data against their own ideals as to what should occur in the classroom. Indeed, as most researchers have experienced at one time or another, the first thing that the average person wants to know about research concerns what has been "proven", what can be "predicted"; in general, does it match the "expectations" which most of us have about research efforts? Rosenshine and Furst (1973) state:

> In place of studies attempting to prove that humane, warm or indirect teachers obtain more student learning than direct teachers, we hope for studies which attempt to improve our understanding of these gross variables (p. 161).

The work of Flanders (1965) is an example of one way in which observational research can improve our under-

standing of the instructional process without being prescriptive. Flanders (1965, p. 8) states his findings regarding the relationship between direct and indirect patterns of teacher influence, and the restriction/ expansion of student action and the increase/decrease in student dependence. There is little doubt that there are those who will read these findings and immediately jump to prescriptive conclusions about one or the other, without realizing that Flanders was pointing out an observed relationship between teacher and student behavior and suggesting that the superior teacher is adept at moving from direct to indirect influence and back again as the situation demands. Smith (1978) supports the use of observational research to improve understanding:

> It is apparent that carefully collected activity ratio data would itself open up a series of interesting questions which the present observation instrument might be useful in investigating. Again these investigations would be concerned with description, relationships, and contingencies, rather than evaluative or prescriptive statements. Thus, the concern would be with accurate, relevant description of the way coaching is, and with studying real effects of planned changes in coaching behavior (p. 16).

Classroom observational research has provided many findings which are potentially useful for those interested in the study of physical education teaching or sports coaching. The reviews by Biddle (1967), Weick (1968), Rosenshine and Furst (1973), Simon and Boyer (1967, 1970, 1974), and others have attempted to draw the threads of

that research together in order to provide guidelines for future investigations, and in the case of Rosenshine and Furst, a "descriptive-correlational-experimental model" for use in observational research.

Observational Analysis and Physical Education

Bookhout (1967) used a modified version of OSCAR, an observational instrument developed by Medley and Mitzel (1959), to determine the pattern of teaching behavior related to the climate of physical education classes. Thirty-six grade 9 girls physical education classes and their respective teachers were employed with each of the classes rated by the pupils on a supportive-defensive continuum. Data on teaching behavior was obtained from four thirty minute observations of each teacher-class unit using OSCAR. The more reliable items of teaching behavior were identified and submitted, along with class climate scores, to factor analysis in order to find teacher behavior patterns. Of the six patterns identified by the analysis, two were considered to be climate related. One pattern, which was named "integrative interactions", was strongly related to supportive climate; the other pattern, "restraining direction", was moderately related to defensive climate. Bookhout argues a case for developing specific instruments for observation in physical education.

A system for observing and describing teacher-student behavior in movement education classes was

developed by Barrett (1969). This instrument provides for the recording of the teacher's verbal behavior as well as the students' movement responses. The events in teaching are recorded under categories within the four major dimensions labelled movement tasks, content, guidance, and student responses. For example, when the teacher presents a movement task for students, the presentation is classified into one of eight categories representing a continuum from command to free exploration. The low levels of inter-observer and intra-observer reliability led Barrett to conclude that her system showed promise, but required further refinement.

A system for recording augmented feedback provided by teachers in physical education was developed by Fishman (1970). The instrument was further refined by Fishman and Anderson (1971) to include twenty specific categories grouped under the following dimensions: form (type of feedback), direction (single student or group), time of delivery of feedback, intent (evaluative, descriptive, comparative, explicative, prescriptive, or affective), general referent (whole or part movement, goal), and specific referent (rate, force and space). The system is quite complex, and as such requires the use of videotape. The recording or coding frequency is dependent upon naturally occurring events of augmented feedback.

Dougherty (1971) reported on the use of the Flanders system of interaction analysis in physical education

settings and suggested two modifications. The first was to categorize meaningful non-verbal activity such as practice on motor skills, which, because of its importance in physical education, should be differentiated from simple silence or confusion; the second was to categorize when the teacher is speaking to an individual rather than the entire group (this is achieved by subscripting the various teacher-talk-categories).

A learning based system to categorize behavior in physical education classes is proposed by Schwartz (1972). The instrument was developed by incorporating aspects of Gentile's (1972) model of skill acquisition which involves the goal, plan, regulatory stimuli, response execution, structuring the environment, feedback, and decision processing. Gentile's learning attributes were operationally defined by Schwartz in terms of teacher behaviors, and the instrument applied to transcripts of the statements of 15 physical education teachers. He concluded that the instrument seemed to provide a satisfactory means of classifying and describing teacher verbal behavior in terms of categories derived from a motor learning model.

Robbins (1973) developed an instrument to analyze teacher verbal behavior in elementary school physical education. It was based upon the Flanders system and on movement pedagogy literature. The 17 category instrument was field tested and deemed by the

investigator to meet reliability and content validity criteria. Robbins concluded that the instrument enabled a reliable and valid analysis and comparison of teacher behaviors in elementary physical education.

Two observation schedules developed by Rushall (1973) for use in physical education settings classify seven teaching behaviors (feedback and rewarding, correcting and prohibiting, questioning, directing, explaining and informing, monitoring and attending, managing and no activity), and nin pupil behaviors (task performance and participation, athlete interaction, questioning, directing, response to teacher, attending, destructive activity, no activity, and injury). The two schedules are designed to be used separately, with observations recorded after each five minutes.

While this writer is somewhat critical of the time sampling methods proposed by Rushall which involve alternation of coding and observing, because potentially important information may be lost, it is also obvious that there could be considerable debate about the arbitrary decisions which are made by system developers concerning the length of the observation interval. A more important criticism of the two instruments presented by Rushall concerns the lack of information provided about the developmental sequence employed in deriving the various categories. For example, the rationale for differentiating the various categories,

particularly the distinctions between "explaining", "feedback", and "correcting", is not discussed by Rushall.

Smith (1974) used an observational system developed by Hough (1967) to examine the consistency between learning environments created by implementing a specialized swimming curriculum, and the critical characteristics that link the theoretical rationale to the curriculum. The Hough system is an adaptation of Flanders' previous work. It includes an additional six categories and was selected by Smith because it included categories well suited to describing activity, and, it was specifically designed to test instructional hypotheses generated from learning theory.

Nygaard (1975) applied the Flanders system of interaction analysis to 40 physical education classes at various grade levels ranging from elementary to "college professional". The primary verbal interaction pattern found by Nygaard involved lecture-followed by silence or confusion-followed by lecture. The utility of this finding is limited by the inappropriateness of using a system designed for analyzing teacher verbal behavior in regular classrooms. Nygaard concluded that the teachers involved in the study were a direct verbal influence in their classes as evidenced in the very low indirect/direct and student talk/teacher talk ratios. This conclusion does not appear to tell us much that is new or useful, however, he did attempt to relate the differ-

ences in the behavior patterns found for men and women to different teaching models and teaching styles.

An observation instrument designed specifically for physical education use by Wilson, Buzzell and Jensen (1975), focuses upon verbal feedback, in particular the type, frequency, and target toward whom the message is directed. The feedback categories include: type (positive performance information, reward, correcting, prohibiting, description), direction (individual, group, or class), and frequency of feedback relative to the events sampled.

The above review provides two important suggestions for those interested in the observational analysis of coaching behavior. The first of these concerns the appropriateness of certain categories derived from classroom observational systems, and the second emphasizes the need for the inclusion of specific categories to cater for the special movement aspects of physical education (and coaching) environments.

Observational Analysis and Coaching

The application of the observational method to the analysis of coaching behavior is even more recent than observational studies of physical education teaching behavior. The two instruments developed by Rushall (1973) for observation of teacher-student behavior in physical education settings are also designed for use in the analysis of coach and athlete behavior.

In a study designed to describe and evaluate the leadership characteristics of minor hockey coaches, Danielson (1974) used Rushall's 7 category instrument to obtain observational measures of coaching behavior in game and practice situations. He found that minor hockey coaches appeared to spend a great deal of time in monitoring-attending behaviors, which strangely enough, Danielson concluded, were unrelated to coaching effectiveness. Intuitively one would expect that monitoring-attending behaviors are related to effective coaching as previously discussed in this chapter. Danielson's conclusion may of course be related in part to inadequacies of the instrument as pointed out in earlier discussion, and partly a function of his interpretation as to what "effective" coaching is, and how this might be measured.

Tharp and Gallimore (1976) applied an observational system they had earlier developed for classroom use, in a study of the coaching behaviors of U.C.L.A.'s John Wooden. They found, as previous observational researchers in physical education found, that the classroom instrument required some modification for use in a movement setting.

> While our standard categories covered most of Wooden's teaching devices, we had to add two new ones to cover what he did. One was the scold-instruct, a criticism followed instantly by how-to-do-it-right; the other could only be called the hustle. Both devices are obvious in any coaching situation, but not in the classroom (p. 75).

The final system contained the following ten categories: instructions, hustles, modelling-positive, modelling-negative, praises, scolds, non-verbal reward, non-verbal punishment, scold/reinstruction, other, and uncodable.

Using an ethnographic approach, interactive behaviors of a swimming coach during training sessions were recorded and described by Gravelle (1977). The salient coaching behaviors extracted from this record included: instructing, commanding, socializing, talks, rides, hustles, confrontations, challenges, modelling-positive, modelling-negative, praising, scolding, non-verbal punishment, ignoring, and attending. These categories of interactive behavior were derived from several hours of daily observation over a five month training season. Gravelle emphasizes that the categories are applicable across coaching environments.

A 20 category observational instrument was developed by Usher (1977) who used the research of Rushall (1973) and Tharp and Gallimore (1976) as a basis. Usher's case study of a nationally recognized hockey coach listed the following as predominant coaching behaviors; directing, attending, monitoring, praise, and explaining. One of the problems with an instrument such as Usher's which contains a relatively large number of categories is that some of them become redundant measures. Smith (1978) is critical of this feature of the Usher schedule:

> His [Usher's] study indicates that three of his categories accounted for almost 69 percent of all the observed behaviors ... However,

> eight categories each accounted for one percent or less, and another five accounted for two percent or less...an imbalance such as in the case of Usher's instrument is to be avoided unless a strong case can be made for the theoretical value of such a large number of infrequently displayed behaviors (p. 10).

Smith, Smoll and Hunt (1977) developed a 12 category coaching behavior assessment system based on social learning theory. There are two main dimensions: reactive behaviors and spontaneous behaviors. Reactive behaviors include responses by the coach to desirable athlete performance, reactions to athlete mistakes, and responses to misbehaviors. Spontaneous behaviors are subdivided into game-related and game-irrevelant behaviors initiated by the coach. The authors report a number of reliability studies involving observer accuracy, stability and inter-rater agreement. However, although the coefficients reported are above 0.88, little detail is given about the statistical analysis, and it would perhaps be expected to achieve acceptable reliability measures with less than 250 behaviors coded on average over an hour and a half. Without additional information little more can be said about the instrument.

The Smith et. al. (1977) instrument was used by Smith, Smoll, Hunt, Curtis and Coppel (1978) in a complex study of the relationships between coaching behaviors and their influence on children participating in youth sports. In phase one of the study the

investigators obtained observational measures of coaching behavior, and through interviews and questionnaires obtained measures of player recall of coaching behavior, players' affective responses to the overall coaching situation, and coaches recall of their own behavior. In phase two the investigators, using social learning principles as a guide, attempted to modify coaching behaviors of a sample of the coaches, and evaluated the success of this intervention with a repeat of the phase one measures. Analysis of the data indicated that the program was highly successful in its effects on coaching behaviors and on the participant's perceptions and attitudes.

An extension of the above research is reported by Smith, Smoll and Curtis (1979) in which Little League Baseball coaches were trained according to the behavioral guidelines previously established (Smith, et. al., 1978). The data showed that the trained coaches differed from controls in both overt and player-perceived behaviors, and that this was accompanied by positive changes in the self esteem of those who played for the trained coaches as compared with those who played for the coaches in the control group. The CBAS (Smith, et. al., 1977) was used to both train the coaches in the experimental group (by providing appropriate models and increasing awareness of coaching behaviors), and to monitor pre-test and post-test behaviors of these coaches.

Both Ritzer (1975), and Rosenshine and Furst (1973), point out that the development of empirical generalizations, the testing of hypotheses, and the development of "theory" are frequently missing in observational research. Brief mention has already been made of the descriptive-correlational-experimental research model proposed by Rosenshine and Furst, and it would appear that the studies by Smith et. al. (1978, 1979) are examples from the observational research on coaching behavior which have gone part of the way toward that proposal. There is a danger, however, in rushing headlong into correlational and experimental research without sufficient time having been spent on developing appropriate measuring instruments which are both reliable and valid.

Hilgard (1964) abstracted six steps from a continuum which he saw starting in "pure" science research and ending in the advocacy and adoption of educational practices. While there may be some debate as to the appropriateness of his early "pure science" steps (Smith, 1975), Hilgard's statement that educational psychologists have tended to spend too little time at the steps which are likely to produce understanding of processes is an important consideration for observational research on coaching.

Smith (1978) has proposed an observational instrument (CBOS) for the analysis of coaching behavior in

practice sessions. It is the product of a variety of sources, including the many years Smith has had as a coach, observer of coaching and teacher; a thorough review and analysis of research in classroom observation, existing teacher behavior observational instruments (Flanders, 1965; Hough, 1967; Robbins, 1973; Simon and Boyer, 1974; Nygaard, 1975; and Anderson, 1976), existing instruments for the analysis of coaching behavior (Smith, Smoll and Hunt, 1977; Tharp and Gallimore, 1976; and Usher, 1977); and discussions with a cross-section of active coaches. Coaching behavior is coded by the observer every five seconds using the ten category definitions and ground rules as a basis for the discrimination of behavior (see Appendix A for a description of the CBOS categories and ground rules). The ten categories include direct and indirect behaviors, and verbal and non-verbal behaviors as indicated in Figure 1.

It would appear that in developing the CBOS, Smith has followed closely the guidelines set down by the various reviewers of observational research such as Rosenshine and Furst (1973), Biddle (1967), and Martin (1976), for the development of categories and ground rules, and for the analysis of data.

Summary

This chapter presents information concerning the types, essential features and uses of observational

Type of Influence	Domain(s) of Behavior	Smith's (CBOS)[a] Categories
Indirect	Nonverbal	1. Monitor
	Verbal-Nonverbal	2. Praise and Reward
Direct	Verbal	3. Instruct
	Nonverbal	4. Demonstration Positive
	Verbal	5. Demonstration Negative
	Verbal	6. Corrective Feedback
	Verbal-Nonverbal	7. Hustles
	Verbal-Nonverbal	8. Scold
	Verbal-Nonverbal	9. Management
Undetermined	Verbal-Nonverbal	10. Other

[a]See Appendix A for a description of the CBOS categories.

FIGURE 1

The Direct-Indirect and Verbal-Nonverbal Dimensions of Smith's Coaching Behavior Observational System.

instruments, and discusses some of the guidelines which have been proposed for the development of observational instruments. Some of the general problems facing developers and users of observational systems are discussed with an emphasis on some of the special requirements necessary for observational instruments intended for use in physical activity settings.

Based on the assumption that coaching is a special form of teaching, some of the research findings and knowledge of relevance to this investigation are briefly discussed, followed by a review of some pertinent research in observational studies of physical education teaching, and sports coaching.

The chapter ends with a brief introduction to the instrument used in this study (CBOS).

CHAPTER III

METHODS AND PROCEDURES

Overview

The primary purposes of this study as set out in Chapter I, are concerned with establishing a data base of videotaped recordings of coaching behaviors, and the coding and analysis of these using the CBOS.

Within these three principal areas, this investigation is concerned with specific objectives. The information which follows in this chapter outlines the methods and procedures employed to collect the data (selection and briefing of subjects, filming equipment and procedures); to code the observed behaviors (details of the CBOS, training procedures, reliability measures); to analyze the coded behaviors (behavior clusters, coaching style indices, sequence matrices, and the computer program); and to arrive at the suggested subdivisions of the Instruct category.

Throughout each of the stages of the study, the investigator was concerned with such issues as the reliability of the measurements, the validity of the various category definitions of the CBOS, and the usefulness of the suggested methods for data analysis. The information below concerning the examination of the development of the CBOS, the reliability measures and training procedures used, and the subdivision of the Instruct category, is related to the reliability and

validity issues. The usefulness of the various forms of data analysis is approached from an experiential point of view. It is here that the experiences of the investigator as recorded in the log maintained during filming, training, coding and analysis, become a valuable aid in making such decisions.

Subjects

The subjects in this investigation were two male team sport coaches at the University of Alberta. The coach of the university men's basketball team was in his third season with the team, and the coach of the university men's volleyball team was in his sixth season with that team.

Both teams were over half way through the 1978-79 inter-university competitive season when videotaping of coaching behaviors during practices was commenced. The completion of the videotaping of the basketball coach coincided with the end of that team's competitive season, while completion of the videotaping of the volleyball coach occurred several weeks before the end of that team's competitive season. In addition, the volleyball team was to take part in an international competitive tour during May, 1979, and hence practice sessions continued for several weeks after the end of the inter-university season.

Each team attended their respective practice sessions on four afternoons of each week when the team travelled

away for weekend competition, and at times attended practice on five afternoons during weeks involving weekend home games.

Practice sessions for the basketball team were generally of 60-75 minutes duration for the period in which videotaping was conducted, while for the volleyball team, practice sessions were for the most part of 110 minutes duration.

There were two assistant coaches working with the basketball coach who were at times in charge of half the team during part of the practice session, allowing the coach to concentrate on smaller groups of players for intensive skills practice. At times during full and half court practices involving all players the two assistant coaches contributed by way of corrective feedback, demonstration, and so on; but generally the coach did most of the instructing, offered feedback and so on, and controlled the practice session.

The volleyball coach employed one assistant coach who participated in a very minor way in the control and conduct of practice sessions. Generally the assistant coach took the role of a player for most, if not all the sessions.

Both coaches were assisted by a team manager who was responsible for tasks involving the supply and maintenance of equipment, first aid, scoring and so on during the practice sessions.

The Coaching Behavior Observational System (CBOS)

The instrument used in this study for the coding and subsequent analysis of coaching behaviors was developed by Smith (1978). The correspondence between the development of the CBOS and the guidelines proposed in the observational research literature has already been briefly discussed in chapter two.

Smith (1978) has described the CBOS as:

> ...a ten category instrument for analyzing the coaching behavior of coaches in a wide variety of different sports. It may also be applied to teachers in most physical education activity settings. The system is designed to provide a systematic approach to the analysis of coaching behavior during practice or workout sessions and is not suitable for studying behavior during games or contests or during lectures or meetings conducted by the coach (p. 1).

The form which the CBOS, as used in this study, took was the result of the careful analysis of the pertinent research literature and considerable attention paid by the developer to the suggested developmental steps. The system was then pilot tested by almost 100 students in an introductory coaching course at the University of Alberta during the 1977-78 academic year. As a result of this pilot testing some modifications were incorporated into the category definitions and ground rules.

Figure 1 in chapter two indicates the main titles of the ten categories and the major dimensions of direct-indirect and verbal-non-verbal behaviors. A brief explanation of each of the ten categories of the CBOS is

included here, while the detailed category definitions and descriptions of the ground rules appear in Appendix A.

Indirect Behaviors

1. Monitor. The coach silently observes individuals or groups as they practice. May include signs of covert thought processes on the part of the coach.
2. Praise and Reward. The coach verbally or non-verbally praises, offers compliments, approval or acceptance of athlete behavior or responses.

Direct Behaviors

3. Instruct. Statements by the coach about what, how, or why to do something. May include questions, requests, and commands.
4. Demonstration, positive. The coach or a participant at the request of the coach, demonstrates how to do something.
5. Demonstration Negative. The coach or a participant at the request of the coach demonstrates an error in performance or something to be avoided.
6. Corrective Feedback. Non-emotional statements by the coach intended to point out errors in performance.
7. Hustles. Statements or actions by the coach intended to activate or intensify previously

instructed behaviors.

8. Scold. Statements or non-verbal activity by the coach intended to scold, criticize, or reject the performance, behavior, feelings or ideas of participants.

9. Management. Verbal or non-verbal activity of the coach associated with housekeeping, management, routine procedures, announcements, and so on.

Undetermined Behaviors

10. Other. Any coaching behaviors which do not fit into the above categories, including statements that cannot be understood and behaviors which are unrelated to the coaching or management tasks at hand.

The above brief explanations are included here to provide continuity as far as the methods and procedures are concerned, in this chapter. It is not intended that the CBOS categories be used to code observed behaviors without reference to the details of the categories and ground rules to be found in Appendix A.

Procedures

The procedures for the study involved a number of phases. After initially developing a familiarity with the CBOS and its developmental background, approval was obtained from the subjects to carry out videotaping in

order to establish a data base. After the data was collected a training phase for the observers was held and various reliability measures conducted. The data was then further coded and analyzed, and finally the procedures for further subdividing the Instruct category of the CBOS were employed.

Data Collection

To facilitate data collection a preliminary briefing was given to the subjects who had offered their assistance with the study. As mentioned in chapter one, in order to reduce observee reactivity it is important to thoroughly explain the rationale for the observation to the subjects before collecting data. At these preliminary discussions the investigator explained the nature of the present study and the possible outcomes, stressing as suggested by Smith (1978), the descriptive rather than evaluative nature of the investigation. This briefing was followed by several visits by the investigator to practice sessions to observe, without recording equipment or the intention of coding behaviors, such things as the available facilities for operating recording equipment, the general structure of the practice sessions, and to familiarize the coach and players with the presence of an observer. It should be noted that both teams and coaches were accustomed to the presence of observers during practice sessions.

Data collection was carried out by videotape recording coaching behaviors with a Sony AVC-3200 video camera equipped with a Sony TV zoom lens, a Sony AV-3600 half-inch monochrome videotape recorder, and a twenty-three inch Electrohome ETV-6 monochrome television receiver. The coach's voice was recorded through a Letrosonics 731 sound system comprised of a Plus Power 90 receiver and extension speaker, and an M-114 Electret microphone wired to an M30r transmitter.

The coach was fitted, prior to the beginning of the practice session, with the transmitting unit. The small microphone was attached to the shirt of the coach and the transmitter was strapped to the coach's belt, around the waist, and in some cases was carried in the coach's hip pocket. The unit was small enough not to be an inconvenience to either coach, a necessary requirement as both coaches were very active throughout the practice sessions.

The receiving set was wired to the videotape recording unit so that the coaches verbal behavior was recorded with the filmed behavior. The extension speaker enabled the investigator to monitor the level of the audio during filming.

The facilities available for setting up equipment varied between the two practice environments in that the

filming position for basketball practices was situated in the spectator seating area some four metres above the midcourt playing surface. This presented some very minor inconvenience in terms of suspended backboards infrequently blocking a clear view of the coach. The best available filming position in the volleyball practice environment was at normal floor level at one end and to the side of the court.

The zoom lens permitted good quality close up and long range filming, an important consideration given the distances involved. Some hints regarding the videotaping of practice sessions appear in Appendix C.

The filming of the coaching behaviors took place over three months with an attempt made to balance the amount of videotaping for each coach. This proved to be a difficult task due to equipment malfunctions, coach illness, and a variety of other related problems. The final tally of observed behaviors showed 4,291 over 230 minutes of observation for the basketball coach, and 7,403 over 370 minutes for the volleyball coach. These tallies were based on the observation of four entire practice sessions for each coach. The figures exclude a small proportion of videotape which was rejected because of poor quality sound or video picture.

Observer Training

The training procedures are outlined here in some detail because of the important relationship between this

phase of the study and the purposes of the investigation as outlined in the first chapter.

The two observers (the investigator and another graduate student who volunteered to assist) memorized the ten category definitions of the CBOS and noted the distinctions referred to in the ground rules and examples presented by Smith (1978). This preliminary learning lasted approximately forty-five minutes and involved constant discussion between the observers, and informal testing of the category numbers, names and definitions.

The observers then worked cooperatively for thirty minutes on the verbal coding by number of the categories of behavior recorded on a five minute segment of videotape selected by the investigator as containing an adequate representation of most, if not all the ten categories of coaching behavior. This selection was made possible as the investigator had previously previewed all the videotapes as a check on the audio and visual quality of each recording. Initially, the observers cooperatively coded the behaviors recorded in the five minute segment, stopping the tape and reviewing as many times as necessary to resolve differences. Constant reference was made to the category definitions and ground rules.

It was found, even at this early training stage, that the coding of certain idiosyncratic behaviors on the part of the two coaches required that some modification be made to the existing ground rules. This is one of

the important justifications for associating ground rules with observation categories. For example, when a training drill was begun by the coach either serving the ball, or passing the ball into play, this was coded as a form of instruction; equivalent to the coach saying "Let's start now". The observers maintained a log of instances where the existing ground rules or category definitions required some modification.

When the two observers were satisfied that they were comfortable with the category definitions, they independently coded a second five minute segment of tape using the prepared data sheet (Appendix B).

The equipment used during the coding phase consisted of a Sony AV-3600 half-inch videotape recorder, a twenty-three inch Electrohome ETV-6 monochrome television receiver, and a Panasonic RV-2400DS dictating cassette recorder. The cassette recorder provided the observers with a recorded "beep" every five seconds to indicate the observation intervals marked on the data coding sheets.

Numbers corresponding to the observed behavior categories were recorded on the coding sheet as they occurred rather than waiting for the end of each five second interval to record the various numbers. This enabled the observers to gain proficiency at coding the data without the need to stop the videotape after each interval.

A second independent coding of the same five minute segment followed after a thirty minute interval, and after a similar rest period a third independent coding was conducted.

At this point intra-observer reliability or stability was calculated across the three samples of coded data. Scott's coefficient (1955) was used in accordance with the guidelines suggested by Flanders (1967) and Hollenbeck (1978). The average coefficients obtained for the two observers were 0.92 and 0.93, which indicated that each observer was independently using the CBOS in a stable or consistent manner.

A small but important point to mention here concerns the need for accurately documenting the exact beginning of each tape and the beginning of any segments used in intra-observer stability, and accuracy checks. In this study, the beginning of each tape was marked and recorded in a log by using distinctive verbal and/or non-verbal coaching behaviors as a guide. In this way it was possible to maintain reasonable accuracy in pin pointing the start of a specific segment.

The next step involved the calculation of inter-observer agreement scores. As the stability coefficients obtained were above the recommended 0.85 acceptable level (Flanders, 1967), it was considered appropriate to calculate the inter-observer agreement scores on the same data. Scott's coefficient was again used, this time,

across observers. The average coefficient of agreement obtained was 0.89. The first calculation of an agreement score resulted in coefficients of 0.86 and 0.82. It was readily apparent that disagreements in the praise category gave rise to these minimally acceptable figures. A recoding of the same segment on the following day, after a preliminary look again at the category definitions, resulted in the acceptable coefficient mentioned above. This result indicated that the two observers were using the CBOS in a similar way in addition to the intra-observer stability as previously established.

The final reliability measure to be calculated before the actual coding of the data began, was a measurement of observer accuracy. The meanings and background to these three reliability measures were discussed in chapter two. The two observers worked co-operatively to code a representative five minute segment of videotaped coaching behaviors of each coach. These segments were coded several times, stopping and reviewing where necessary, until perfect agreement was reached. In this way an accuracy criterion was established for each coach for subsequent accuracy checks. These accuracy criteria were deemed to be an accurate account of each subject's coaching behavior during that specific segment, given the limitations imposed by the use of the CBOS and the conditions of observation.

In summary, three measures of reliability were

established prior to the coding of the data used in the study. These measures indicated that each observer was coding in a consistent manner, and that the two observers were coding the same material in a similar way. The third measure established, by consensual agreement, was two accuracy criteria against which subsequent coding performance was to be compared.

Coding the Data

All of the data was coded in seven days following the training and reliability checks described above. Because of the relatively short duration of this coding period as compared to that described in the reliability literature (Johnson and Bolstad, 1973; Kazdin, 1977), the investigator felt that it would be necessary to calculate the three measures discussed above again, and this was done after approximately seventy percent of the data had been coded.

Because only the investigator was actively involved in the coding of data there was little point in calculating inter-observer agreement again. One could not expect a non-participating observer to retain the five distinctions and coding skills established during earlier training. An informal check of inter-observer agreement produced coefficients of 0.70 and 0.62 which lent support to the deletion of this measure.

The obtained coefficients for the participating observer were all above the level of general acceptance;

the intra-observer coefficient as measured by Scott's method was 0.86 for the basketball coach and 0.94 for the volleyball coach. The accuracy coefficient as judged against the criterion previously established was 0.84 for the basketball coach and 0.97 for the volleyball coach. The lower figure for the basketball coach was in part attributed to problems with the videotape phasing, that is, the starting points were not accurately aligned.

Following these checks on stability and accuracy, the remainder of the coding of the data was completed.

It should be mentioned at this point that the videotapes were coded in a random order to reduce the effects of observer drift. This precautionary measure was a simple one to make as suggested by Kazdin (1977), and when supported by the accuracy coefficient indicates an absence of bias in the coding of the data, both across subjects and practice sessions.

Data Analysis

The data which had been coded onto the prepared sheets was transferred onto IBM data cards. These cards were analyzed by the IBM 360/67 computer facilities at the University of Alberta. The program used for the analysis was developed and documented by Burnett, Flatham and Westrom of the Division of Educational Research Services (DERS), and titled "Flanders Interaction Analysis, Test 13".

The CBOS may be viewed as a modification of the Flanders Interaction Analysis System (1965), and hence this program was an available and convenient approach to data analysis. The program documentation provided by DERS gives the following description:

> Input is a stream of cell indicators. The program uses these to construct an interaction matrix and then outputs the matrix in terms of percentage cell frequencies and row and column totals. Automatic comparisons are made between all such matrices input, and the user may ask for special comparisons. The user may also request special ratios to be produced for each matrix (p. 1).

The data treatments proposed by Smith (1978, pp. 19-23) are all conveniently handled by this program. Basically, there are two transformations of the originally coded category numbers which provide in the first instance, information concerning the behavior clusters, and coaching style indices; and in the second instance, information about the sequential aspects of the observed behaviors.

The first level of analysis provides category frequencies and percentages which are used to develop the various clusters and indices. These describe various aspects of coaching behavior which may be important in following fluctuations and describing variations between coaches of different levels and sports (Smith, 1978). It is, however, important to note as Smith points out:

> Such measures should be considered in the context of the description of the physical setting, the activity time percentages, and

> ratio, and the explanatory comments that appear...(1978, p. 19).

Behavior clusters. These groupings of categories include: *Primary coaching behavior* (C) which is calculated simply by adding the totals for categories 3, 4, 5, and 6. It describes behaviors which include: "...most, if not all, of the communication regarding cognitive and motor aspects of performance" (Smith, 1978, p. 19). *Affective behavior* (A) is the sum of behaviors observed in categories, 2, 7, and 8, and concerning behaviors which are: "...primarily, but not exclusively, concerned with emotional aspects or affect" (Smith, 1978, p. 21). *Indirect behavior* (I) defines the sum of behaviors observed in categories 1 and 2. Smith (1978, p. 21) states: "It connotes an observing, encouraging, supportive, style of teaching or coaching...". *Direct behavior* (D) is related to an authoritative coaching style and is calculated by adding the totals for categories 3 through 9. *Positive behavior* (P) is calculated by adding the totals for categories 2 and 4, while *Negative behavior* (N) is the sum of categories 5 and 8.

Coaching style indices. These provide an index of what Smith refers to as "the coaching climate", which may be considered as analogous to Flanders' (1965, 1970) "teaching climate". The various indices are calculated using specific behavior clusters as denominators and numerators of the ratio. The *Primary coaching/affective*,

ratio (C/A), provides a ratio of communication to affective behaviors. *The Indirect/direct, ratio* (I/D) has been important in the work of Flanders (1965) and others (Amidon and Hough, 1967) who have followed his approach. Smith (1974) also found this ratio useful as an index of shifting physical education teaching style. The *Primary coaching/monitor, ratio* (C/M) provides information about the proportion of "active" to "inactive" coaching. The term inactive as used here has no evaluative connotations (the importance of monitoring has been previously discussed). The *Positive/negative, ratio* (P/N) provides an index of the mix between the positive and negative behaviors of coaches, a measure which Smith et. al. (1978b) found to be quite useful.

A further coaching style index which Smith does not include, was included in this study. Its importance became apparent during the coding of the videotapes as a means of describing some of the observed behavioral variance. The index was called the *Primary coaching/manage, ratio* (C/Ma) and is calculated by adding the totals of categories 3 through 6 and dividing this sum by the total for category 9.

Behavioral matrices. The second level of raw data transformation concerns the construction of behavioral matrices, which have also been developed as an outcome of the observational research on classroom teaching (Flanders, 1965). These matrices preserve the sequential

aspects of the observed behavior and enable the investigator to consider at least three features of this; steady state behaviors which continue for more than five seconds, antecedent behaviors, and following behaviors. Smith (1978) states that sequential information of this type would be:

> ...useful in studying the coaching role, in studying particular coaches, in identifying real relationships among behaviors, and in planning and assessing planned changes in coaching behavior (p. 28).

Smith also suggests that it may prove worthwhile to consider particular areas of matrices in the same way as proposed by Flanders. For example, the relationship between direct and indirect behaviors, and the nature of the primary coaching behaviors, can be assessed by focusing on particular matrix areas.

Data analysis of the type proposed by Smith can be conducted manually, or with the aid of an appropriate computer program. An illustrative example of the calculation of the various clusters and ratios, and the development of a behavioral matrix from a sample of the raw data of this investigation, is presented in Appendix B.

The computer program provided the necessary sequence matrices and category frequencies and percentages which enabled the investigator to calculate the required clusters and ratios by visual inspection of the obtained outputs. The steady state, preceding and following behaviors were also calculated by the investigator in

this way.

The computer analysis provided matrices relevant to three different levels of the obtained data. Individual practice session matrices were constructed for each coach, a combined matrix for each coach was constructed based on a combination of the respective practice sessions, and finally, a total matrix was constructed which combined the practice session data for both coaches.

Subdivision of the Instruct Category

Although Smith (1978, pp. 39-47) has outlined the correspondence between his development of the CBOS and the guidelines suggested by some of the reviewers of observational research (Biddle, 1967; Adams, 1972; Simon and Boyer, 1974; Martin, 1976), he points to the need for further examination of the "...relevant generic sub-blocks of [instruction]" (p. 32). He suggests that this will lead to a "...higher power of resolution in examining the complex behaviors subsumed by this major block of coaching behavior" (p. 32).

To supplement the investigator's own experience with teaching and coaching instruction over the last fifteen years, an attempt was made to examine some definitions of instruction in order to arrive at some preliminary titles for a number of "generic sub-blocks". However, after consulting the Dictionary of Education, The Encyclopaedia of Education, and Webster's Dictionary, little was

achieved in terms of preliminary titles.

The pedagogical literature concerned with the instructional process also failed to provide a useful definition. Perhaps it is the case that the term "instruction" is so commonly used that it is felt unnecessary to define it. Bruner (1966) for example, while concerned with the important question of the development of a theory of instruction, does not attempt to define instruction. He emphasizes the importance of establishing a predisposition in the learner toward learning, an optimal structure for a body of knowledge, an effective progressive sequence of instruction, and the nature and pacing of rewards and punishments. Although he does not provide a useful definition, his illustrative examples did assist with the formulation of the category titles which follow.

The observational research literature, for example Simon and Boyer (1974), was also useful as a source of various approaches toward instructional categories. Flanders (1965), Hough (1967) and Wankel and Pabich (1977), among others, provide useful illustrative examples of different types of instruction.

An examination of specific observational instruments (the Flanders FIAS, Hough's OSIA and Smith's CBOS) also provided an indication of some potential sub-categories of instruction.

The log maintained by the investigator during the

observer training and data coding stages of this study was an additional source of information.

The investigator developed a list of sub-categories on the basis of the sources of information above, and this list of titles, descriptions and examples was submitted to the scrutiny of several experienced teacher/coaches, which included two members of the Physical Education Faculty at the University of Alberta, and two graduate students, who all had a minimum of ten years experience as teacher/coaches. The list of sub-categories was independently presented to each person with the request that each sub-category be examined in the light of their own experience in the field, and that suggestions be made regarding the mutual exclusiveness, possible omissions and modifications of the list.

As a result of these discussions, the original list (Appendix D) was modified and is presented as follows:

<u>Command, Order, Direct</u>

Includes both verbal and non-verbal directions, orders and commands to which compliance is expected. Examples of verbal commands: "Stop! Run in here!" "Red team on offense, blue on defence!" Examples of non-verbal commands include actions or gestures by the coach indicating for example, the direction in which players are to run. Non-verbal commands also include instances where the coach puts the ball into play to begin a drill, and blowing a whistle to stop activity.

Lecture, Explain, Expand

Includes verbally giving facts or opinions about content or procedure. Examples: "In this drill it is important that this player runs right through to the baseline so that he can exchange under the basket". "The key to our defense is the blocking power of our three front line players. So, when you are on the net you must be quick to read the offensive moves of the other team".

Request

Includes verbal statements by the coach which are not authoritative as in commands, and which do not require compliance. The receiver is given the option to reject or accept the statement. These statements are often made quietly, person to person. Examples: "Now you look as though you are ready to try...". After a player answers to a previous question - "Well let's try this...".

Question, Implied Question

Includes verbal and non-verbal questions to which answers or expected. These questions usually concern content or procedure. Example of a verbal question: "How could you counter a move like that?" Example of a non-verbal or implied question: an extended pause with raised eyebrows by a coach in an attempt to turn a player question back to the player for the answer.

Response to Player Question

Includes both verbal and non-verbal direct answers

to player questions. Such verbal answers may give information or opinion, but must be responses that answer or are directed toward answering player questions. Example: "No, you need to remain close to the baseline." Non-verbal responses to player questions include shaking the head to indicate a "yes" or "no" or a gesture indicating a "don't know".

Confirmation Feedback

The coach indicates to the athlete, either verbally or non-verbally, that his performance or statement is correct. This confirmation and acceptance of ideas and performance is affect free. The feedback may indicate that performance was satisfactory and should be repeated in the same way, or that although satisfactory, performance will be improved further by specified modifications. Verbal example: "OK, your hand contacted the ball in just the right position." A non-verbal example may include nodding the head to indicate to a player that his performance is acceptable.

Content Analysis

The list of instructional sub-categories was used in the content analysis of transcribed samples of each coach's recorded behaviors. The choice of the samples was based on behaviors which had been previously coded as extended periods of instruction (category 3). The explanatory comments contained in the log maintained during filming and coding

were an additional guide to the selection of the samples.

CHAPTER IV

RESULTS

CBOS Behavioral Analysis

The descriptive results of the CBOS analysis of coaching behaviors during practice sessions are presented in the following Tables. The category counts and percentages, behavior clusters and coaching style indices are presented first, followed by the sequential aspects of the observed behaviors.

The results are presented at these two levels of analysis separately for each coach, with an example of a single practice session and combined practice sessions provided. In addition, the data from all practice sessions for both coaches is combined for both levels of analysis. An analysis of the results of each sample appears in Appendix E.

The single practice sessions for each coach were selected as being representative of the behaviors observed during this study, the selection being based on a thorough review of all the obtained data. To enhance the reader's understanding of the results presented in these single practice session examples, a brief description of the conditions specific to the collection of that data are provided.

To maintain anonymity the coaches are here referred to as Coach A and Coach B.

TABLE 1

CBOS Analysis of Sample 4, Coach A.

CBOS Category Percentages

Category	Count	Percent
1. Monitor	415	40.0
2. Praise	108	10.4
3. Instruct	319	30.8
4. Demo +	4	0.4
5. Demo -	1	0.1
6. Correct	49	4.7
7. Hustle	16	1.5
8. Scold	12	1.2
9. Manage	111	10.7
10. Other	2	0.2
TOTALS	1037	100

Total practice time = 60 minutes
Observations/minute = 17.28

Behavior Cluster Measures*

1. Primary Coaching (C) categories 3-6 (373/1037) 100 = 36.0%
2. Affective Coaching (A) categories 2, 7, 8 (136/1037) 100 = 13.1%
3. Indirect Coaching (I) categories 1, 2 (523/1037) 100 = 50.4%
4. Direct Coaching (D) Categories 3-9 (512/1037) 100 = 49.4%
5. Positive Coaching (P) Categories 2, 4 (112/1037) 100 = 10.8%
6. Negative Coaching (N) Categories 5, 8 (13/1037) 100 = 1.3%

*Because the categories involved in the various clusters overlap, the total percent figure exceeds 100.

Coaching Style Indices'

1. Primary Coaching/affective C/A = 373/136 = 2.74
2. Primary Coaching/monitor C/M = 373/415 = 0.90
3. Primary coaching/manage C/Ma = 373/111 = 3.36
4. Indirect/direct I/D = 523/512 = 1.02
5. Positive/negative P/N = 112/13 = 8.62

TABLE 2

CBOS Behavioral Matrix, Sample 4, Coach A

	1 MONITOR	2 PRAISE	3 INSTRUCT	4 DEMO +	5 DEMO -	6 CORRECT	7 HUSTLE	8 SCOLD	9 MANAGE	10 OTHER	TOTAL
Monitor 1	158	71	150	0	0	16	5	4	11	0	415
Praise 2	33	7	46	0	0	8	1	0	13	0	108
Instruct 3	175	19	75	3	0	8	9	1	28	1	319
Demo + 4	0	0	3	0	0	1	0	0	0	0	4
Demo - 5	0	0	0	0	0	1	0	0	0	0	1
Correct 6	12	2	17	1	1	1	0	2	3	0	49
Hustle 7	7	0	6	0	0	1	1	0	1	0	16
Scold 8	2	0	2	0	0	2	0	5	1	0	12
Manage 9	27	9	19	0	0	1	0	0	54	1	111
Other 10	1	0	1	0	0	0	0	0	0	0	2
TOTAL	415	108	319	4	1	49	16	12	111	2	1037
%	40.0	10.4	30.8	0.4	0.1	4.7	1.5	1.2	10.7	0.2	100

Steady state cells along the diagonal indicate the behaviors that continued for more than five seconds. For example, instruction continued for more than five seconds on 75 of the 319 occasions it was recorded.

Preceding behaviors. To see which behaviors preceded behavior in any given category, go to the *column* of the category and examine each cell in the column. For example, of the 319 instances of instruction, 150 were preceded by monitoring, and 46 were preceded by praise.

Following behaviors. Go to the *row* of a given category and examine each cell. For example, of the 319 instances of instruction, 175 were followed by monitoring, and 28 by management behaviors.

TABLE 3

CBOS Analysis of Combined Practice Sessions, Coach A.

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	1777	41.4
2. Praise	504	11.7
3. Instruct	1489	34.7
4. Demo +	26	0.6
5. Demo -	2	0.0
6. Correct	204	4.8
7. Hustle	103	2.4
8. Scold	20	0.5
9. Manage	158	3.7
10. Other	8	0.2
TOTALS	4291	100

Total practice time = 231 minutes
Observations/minute = 18.6

Behavior Cluster Measures*

1. Primary Coaching (C) categories 3-6 (1721/4291) 100 = 40.1%
2. Affective Coaching (A) categories 2, 7, 8 (627/4291) 100 = 14.6%
3. Indirect Coaching (I) categories 1, 2 (2281/4291) 100 = 53.2%
4. Direct Coaching (D) categories 3-9 (2002/4291) 100 = 46.7%
5. Positive Coaching (P) categories 2, 4 (530/4291) 100 = 12.35%
6. Negative Coaching (N) categories 5, 8 (22/4291) 100 = 0.5%

*Because the categories in the clusters overlap the total percent figure exceeds 100.

Coaching Style Indices

1. Primary coaching/affective C/A = 1721/627 = 2.75
2. Primary coaching/monitor C/M = 1721/1777 = 0.97
3. Primary coaching/manage C/Ma = 1721/158 = 10.89
4. Indirect/direct I/D = 2281/2002 = 1.14
5. Positive/negative P/N = 530/22 = 24.09

TABLE 4

CBOS Behavioral Matrix, Combined Samples, Coach A.

	MONITOR 1	PRAISE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor 1	692	285	665	0	0	75	35	8	16	1	1777
Praise 2	159	40	244	0	0	27	10	2	21	1	504
Instruct 3	766	151	413	17	0	46	51	2	41	2	1489
Demo + 4	2	0	17	3	1	3	0	0	0	0	26
Demo - 5	1	0	0	0	0	1	0	0	0	0	2
Correct 6	60	12	76	5	1	41	2	2	5	0	204
Hustle 7	51	4	35	0	0	4	5	1	2	1	103
Scold 8	3	1	6	0	0	4	0	5	1	0	20
Manage 9	41	10	30	0	0	2	0	0	72	3	158
Other 10	2	1	3	1	0	1	0	0	0	0	8
TOTAL	1777	504	1489	26	2	204	103	20	158	8	4291
%	41.4	11.7	34.7	0.6	0.0	4.8	2.4	0.5	3.7	0.2	100

Steady state cells along the diagonal indicate the behaviors that continued for more than five seconds. For example, monitoring occurred for more than five seconds on 692 of the 1777 occasions it was recorded.

Preceding behaviors. Go to the *column* of the category and examine each cell. For example, praise was preceded by monitor on 285 of the 504 occasions it was recorded.

Following behaviors. Go to the *row* of the category and examine each cell. For example, of the 1489 instances of instruction, 766 were followed by monitoring, and 151 by praise.

TABLE 5

CBOS Analysis of Sample 3, Coach B

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	755	36.5
2. Praise	189	9.1
3. Instruct	831	40.1
4. Demo +	1	0.0
5. Demo -	1	0.0
6. Correct	32	1.5
7. Hustle	28	1.4
8. Scold	5	0.2
9. Manage	224	10.8
10. Other	4	0.2
TOTALS	2070	100

Total practice time = 105 minutes
Observations/minute = 19.7

Behavior Cluster Measures*

1. Primary Coaching (C) categories 3-6 (865/2070) 100 = 41.8%
2. Affective Coaching (A) categories 2, 7, 8 (222/2070) 100 = 10.7%
3. Indirect Coaching (I) categories 1, 2 (944/2070) 100 = 45.6%
4. Direct Coaching (D) categories 3-9 (1122/2070) 100 = 54.2%
5. Positive Coaching (P) categories 2, 4 (190/2070) 100 = 0.9%
6. Negative Coaching (N) categories 5, 8 (6/2070) 100 = 0.3%

*Because the categories in the clusters overlap, the total percent figure exceeds 100.

Coaching Style Indices

1. Primary coaching/affective C/A = 865/222 = 3.90
2. Primary coaching/monitor C/M = 865/755 = 1.15
3. Primary coaching/manage C/Ma = 865/224 = 3.86
4. Indirect/direct I/D = 944/1122 = 0.84
5. Positive/negative P/N = 190/6 = 31.67

TABLE 6

CBOS Behavioral Matrix, Sample 3, Coach B.

		MONITOR 1	PRAISE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	99	172	421	0	0	24	18	4	15	2	755
Praise	2	13	4	166	0	0	2	0	1	2	1	189
Instruct	3	629	9	172	1	0	2	7	0	11	0	831
Demo +	4	0	0	0	0	1	0	0	0	0	0	1
Demo -	5	0	0	1	0	0	0	0	0	0	0	1
Correct	6	1	0	27	0	0	4	0	0	0	0	32
Hustle	7	1	3	21	0	0	0	2	0	1	0	28
Scold	8	0	0	4	0	0	0	1	0	0	0	5
Manage	9	12	1	16	0	0	0	0	0	194	1	224
Other	10	0	0	3	0	0	0	0	0	1	0	4
TOTAL		755	189	831	1	1	32	28	5	224	4	2070
%		36.5	9.1	40.1	0.0	0.0	1.5	1.4	0.2	10.8	0.2	100

Steady state cells along the diagonal indicate the behaviors that continued for more than five seconds. For example, category 9 Manage continued for more than 5 seconds on 194 of the 224 instances which were recorded.

Preceding behaviors. Go to the *column* of a given category and examine each cell. For example, of the 189 instances of praise, 172 were preceded by monitoring.

Following behaviors. Go to the *row* of a given category and examine each cell. For example, of the 831 instances of instruction, 629 were followed by monitoring.

TABLE 7

CBOS Analysis, Combined Practice Sessions, Coach B.

CBOS Category Percentages

Category	Count	Percent
1. Monitor	2704	36.5
2. Praise	603	8.1
3. Instruct	2935	39.6
4. Demo +	5	0.1
5. Demo -	4	0.1
6. Correct	116	1.6
7. Hustle	287	3.9
8. Scold	41	0.6
9. Manage	658	8.9
10. Other	50	0.7
TOTALS	7403	100

Total practice time = 371 minutes
Observations/minute = 19.95

Behavior Cluster Measures*

1. Primary Coaching (C) categories 3-6 (3060/7403) 100 = 41.3%
2. Affective Coaching (A) categories 2, 7, 8 (931/7403) 100 = 12.6%
3. Indirect Coaching (I) categories 1, 2 (3307/7403) 100 = 44.7%
4. Direct Coaching (D) categories 3-9 (4046/7403) 100 = 54.7%
5. Positive Coaching (P) categories 2, 4 (608/7403) 100 = 8.2%
6. Negative Coaching (N) categories 5, 8 (45/7403) 100 = 0.6%

*Because the categories in the clusters overlap, the total percent figure exceeds 100.

Coaching Style Indices

1. Primary coaching/affective
   C/A = 3060/931 = 3.29
2. Primary coaching/monitor
   C/M = 3060/2704 = 1.13
3. Primary coaching/manage
   C/Ma = 3060/658 = 4.65
4. Indirect/direct
   I/D = 3307/4046 = 0.82
5. Positive/negative
   P/N = 608/45 = 13.51

TABLE 8

CBOS Behavioral Matrix, Combined Practice Sessions, Coach B.

		MONITOR 1	PRAISE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	524	391	1526	1	1	69	119	18	41	14	2704
Praise	2	101	16	451	0	0	4	19	1	8	2	602
Instruct	3	1932	145	652	1	0	22	128	11	41	3	2935
Demo +	4	0	0	4	0	1	0	0	0	0	0	5
Demo -	5	0	0	1	2	0	1	0	0	0	0	4
Correct	6	12	2	83	1	2	12	2	1	0	1	116
Hustle	7	81	41	133	0	0	8	14	7	2	1	287
Scold	8	6	2	25	0	0	0	5	3	0	0	41
Manage	9	40	6	47	0	0	0	0	0	550	14	657
Other	10	8	0	13	0	0	0	0	0	16	15	52
TOTAL		2704	602	2935	5	4	116	287	41	657	52	7403
%		36.5	8.1	39.6	0.1	0.1	1.6	3.9	0.6	8.9	0.7	100

Steady state cells along the diagonal indicate the behaviors that continued for more than five seconds. For example, of the 657 instances of management, 550 lasted for more than five seconds.

Preceding behaviors. Go to the *column* of a given category and examine each cell. For example, of the 602 instances of praise, 391 were preceded by monitoring, and 145 by instruction.

Following behaviors. Go to the *row* of a given category and examine each cell. For example, of the 287 instances of hustles, 133 were followed by instruction, and 81 by monitoring

TABLE 9

CBOS Analysis, Coaches A and B Combined Over All Practice Sessions.

CBOS Category Percentages

Category	Count	Percent
1. Monitor	4481	38.3
2. Praise	1107	9.5
3. Instruct	4424	37.8
4. Demo +	31	0.3
5. Demo -	6	0.1
6. Correct	320	2.7
7. Hustle	390	3.3
8. Scold	61	0.5
9. Manage	816	7.0
10. Other	58	0.5
TOTALS	11694	100

Total Practice time = 602 minutes
Observations/minute = 19.43

Behavior Cluster Measures*

1. Primary Coaching (C) categories 3-6 (4781/11694) 100 = 40.9%

2. Affective Coaching (A) categories 2, 7, 8 (1558/11694) 100 = 13.3%

3. Indirect Coaching (I) categories 1, 2 (5588/11694) 100 = 47.8%

4. Direct Coaching (D) categories 3-9 (6048/11694) 100 = 51.7%

5. Positive Coaching (P) categories 2, 4 (1138/11694) 100 = 9.7%

6. Negative Coaching (N) categories 5, 8 (67/11694) 100 = 0.6%

*Because the categories in the clusters overlap, the total percent figure exceeds 100.

Coaching Style Indices

1. Primary coaching/affective C/A = 4781/1558 = 3.07

2. Primary coaching/monitor C/M = 4781/4481 = 1.07

3. Primary coaching/manage C/Ma = 4781/816 = 5.86

4. Indirect/direct I/D = 5588/6048 = 0.92

5. Postive/negative P/N = 1138/67 = 16.99

TABLE 10

CBOS Behavioral Matrix, Coaches A and B Combined over all Practice Sessions

		MONITOR 1	PRAISE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	1216	676	2191	1	1	144	154	26	57	15	4481
Praise	2	260	56	695	0	0	31	29	3	29	3	1106
Instruct	3	2698	296	1065	18	0	68	179	13	82	5	4424
Demo +	4	2	0	21	13	2	3	0	0	0	0	31
Demo -	5	1	0	1	2	0	2	0	0	0	0	6
Correct	6	72	14	159	6	3	53	4	3	5	1	320
Hustle	7	132	45	168	0	0	12	19	8	4	2	390
Scold	8	9	3	31	0	0	4	5	8	1	0	61
Manage	9	81	16	77	0	0	2	0	0	622	17	815
Other	10	10	1	16	1	0	1	0	0	16	15	60
TOTAL		4481	1106	4424	31	6	320	390	61	815	60	11694
%		38.3	9.5	37.8	0.3	0.1	2.7	3.3	0.5	7.0	0.5	100

Steady state cells along the diagonal indicate the behaviors that continued for more than five seconds. For example, of the 4424 instances of instruction, 1065 continued for more than five seconds.

Preceding behaviors. Go to the *column* of a given category and examine each cell. For example, of the 1106 instances of praise, 676 were preceded by monitoring, and 296 by instruction.

Following behaviors. Go to the *row* of a given category and examine each cell. For example, of the 61 instances of scolds, 31 were followed by instruction. For the 390 instances of hustles, 168 were followed by instruction, and 132 by monitoring.

Subdivision of the Instruct Category

The videotaped verbal behaviors of the coaches during the selected samples were transcribed and content analyzed according to the following sub-categories described previously in chapter three:

1. Command, order, direct.
2. Lecture, explain, expand.
3. Request.
4. Question, implied question.
5. Response to player question.
6. Confirmation feedback.

An instructional sequence from the recorded behaviors of each coach is presented below with the various statements coded by the number representing the appropriate sub-category. Behaviors coded according to the other nine categories of the CBOS are indicated in parentheses.

Sample 1, Coach A

Instructional sub-category	Coaching Behavior
1	OK, post men down here, point and wings down there. (management activity)
1	OK, let's have Ash and Goad come with me, the other three stay with the coach. (management activity)
1	OK let's go on handball.
2	Just a sec, I've got to set it up.(player question)
4	Eh? (player repeats question)

5 Yeh
(management)

2 It's a game of handball off the toss bench. You gotta catch it then get out of the way and let the other guy catch it. The ball's gotta stay within the white boundaries. If you miss the ball you lose a point. So pass it hard or soft, whatever, then get out of the way so the other guy can get it.

(player question)

5 Yeh.
(player question)

5 Back
(player question)

5 No, make it this white line.
(player makes statement)

6 Yeh.

1 Coach hands ball to players - OK, go ahead.
(players begin game, coach monitors)

6 OK, that's one.
(management)
(hustle)

6 Two, that's two nothing.

1 Hold it, fix this thing, hold it.
(management)

6 (a series of point scores given by coach)
(coach gives corrective feedback to a player, followed by -)

3 Mix it up, throw it hard soft -

<u>Sample 2, Coach B</u>

Instructional sub-category	Coaching Behavior
1	OK fellas.
3	Let's have everybody down at the far end line, eh.
1	Let's go, go.
2	Do it in two laps, eh. Two groups. We're gonna have, (pause) two groups of four, eh. (management - counts players)
2	Let's have Gordie, Ronnie and Al. (player comment)
6	OK, and Jimmy.
4	OK?
1	OK, we'll start off. We're gonna use this court and the other court.
4	OK?
2	Ten and ten, sprint to the net, coast to the end line, go over there, and the same thing. We'll do one thing at this net, one thing at the other net, then we'll change each time we're down here. (player question)
5	No.
1	Ready, Go! (hustle)
2	Sprint to the net, coast to the end line.
1	Go! (monitors)
2	Same thing down here.
1	Go. (monitor, hustle)

2	OK, ten and ten, 3 steps 360, 4 steps 360, sprint to the end line.
1	Go!
2	We're starting off a little easier today.
1	Sprint to the end line - Go! (player activity)
1	OK fellas, balls over here.
2	Let's get our hands back in shape now. (players return balls to basket)
2	OK, we're just gonna do some of the things we did last week a bit, for a bit.
4	OK, ready?
1	Coach sets ball for player to hit.
2	Tell me if it's OK for the lights there.
4	Is it OK?
1	OK, let's call it.
1	Coach sets ball for each player.

CHAPTER V

DISCUSSION

This chapter presents a discussion of results pertaining to videotape records of actual coaching behaviors validation of the CBOS, the descriptive statistics for the CBOS categories, clusters, indices and sequence matrices.

The Data Base

The relative infancy of the use of the observational method in the study of coaching behavior points to the need for the development of a data base of recorded observations, and for a variety of samples of such observations. An improved understanding of the coaching process ultimately requires information from research conducted at various levels of coaching involvement, from the volunteer community coach to the international and professional coach, and from the coach of youth sports through to the coach of the elderly participant. We need to have information obtained from the structured observation of coaches from the many available team and individual sports. We need to be able to chart variations in coaching behavior across an entire season, and at the micro-level, at various stages leading up to and following weekly competition.

Because of the complexity of the coaching process, as is the case with any investigation of human behavior, there is a need for a variety of research approaches. Smith and Smoll (1978) emphasize this in their discussion

of research approaches in youth sports. They also propose that observational research has an important part to play in providing information about aspects of the complex sporting phenomenon.

This observational study is preliminary in the sense that it is primarily concerned with investigating the usefulness of a recently available instrument designed for the reliable, structured observation of coaching behavior in practice settings. Because of this emphasis, the data collected provides little more than the beginning of a data base which may eventually be expanded to include information pertinent to the issues mentioned above.

The example of a suggested format for a coding sheet presented in Appendix F represents an attempt to begin to develop procedures for establishing data banks that are accessible to other researchers who may be interested in subsequent cross-validation of previously obtained research findings (Rosenshine and Furst, 1973).

In developing data collection procedures in this study, there were some guidelines available (Smith, 1978), but there was also a need for decisions which might best be described as exploratory. This was also evident in developing suggestions for the subdivision of the Instruct category. Flanders (1965) suggests that a minimum of twenty minutes of "homogenous" activity be recorded in classroom observational research. There is little additional information available in the literature to guide the amount of data

recording. In this study, the investigator chose to record coaching behaviors observed during *entire* practice sessions in an attempt to provide a wide range of coaching behaviors to which the CBOS could be applied. Later studies which focus upon particular aspects of coaching behavior will be able to use this same data base, but limit the analysis to a restricted range of the observed behaviors, such as for example, instructional behaviors.

In accordance with the suggestions provided by Flanders (1965) and Anderson (1972) it can be reasonably assumed that the 11,694 behaviors coded in this study are representative of the normal practice session behaviors of each coach.

Validation of the CBOS

Implicit in the concept of validity as previously discussed is the mutual exclusiveness of the categories and the range of behaviors account for by them.

The mutual exclusiveness of the CBOS categories is partly reflected in the high accuracy and stability coefficients obtained during the observer training and data coding stages of this study. The accuracy coefficient averaged 0.90 for the coding stage, while the obtained stability coefficient averaged 0.92 during training, and 0.90 during data coding. These exceed the 0.85 level suggested as acceptable by Flanders (1965) and Hollenbeck (1978). Appendix G shows the calculation of Scott's coefficient.

The capacity of the CBOS categories to satisfactorily describe the range of observed behaviors is also reflected to some extent in the coefficients mentioned above. Obviously, if an observer encounters behaviors which are not described by the categories, and there is some confusion as to where such behaviors belong, the possibility of obtaining acceptable stability and accuracy coefficients must be limited.

During both observer training and data coding it was, however, apparent that certain, often idiosyncratic, behaviors did not fit the available CBOS category definitions and ground rules. The following suggestions are made in this regard:

1. Where the coach is involved in activity as a player this should be coded as management, although clearly he or she may exhibit other behaviors which must be appropriately coded.
2. Where a coach is involved in his or her own warm up or stretching activities, this should also be coded as management.
3. In situations where the coach calls out the score it may be coded as instruction or a scold depending on the context, emphasis and way it is received by the player(s).
4. Hitting, serving or passing the ball into play to begin a drill is a form of instruction.
5. When the coach refers to prepared notes this should

be coded as management.

6. An apology by the coach for inadequate placement of a ball or a mistake should be interpreted as: "It's not your fault, it's mine". This is a form of confirmatory feedback and should be scored as instruction.

A "valid" instrument is a necessary, but not sufficient requirement for accurate and consistent coding of observed behavior. An observer who for one reason or another fluctuates in his or her capacity to accurately and consistently code observed behaviors, may do so despite the previously established validity of the instrument. Therefore, inter-observer agreement measures should be combined with accuracy and consistency tests as they were in this study.

CBOS Descriptive Analyses

A purpose of this study was to follow the suggested guidelines proposed by Smith (1978) for the analysis of data, and to suggest modifications where necessary. This latter requirement is a difficult task for a preliminary study such as this one, however, the discussion which follows does contain certain proposals for additions to the previously established guidelines.

To recapitulate the analyses conducted in the study, the recordings of observed coaching behaviors were coded according to the CBOS categories and associated ground rules, and analyzed to provide percentage and frequency counts for the categories. These were grouped together

in various behavior clusters, which in turn were used in the calculation of a number of coaching style indices. A second type of analysis resulted in behavioral matrices which preserve sequential aspects of the observed behaviors.

An understanding of the different categories of the CBOS or similar instruments may itself be a means of increasing the awareness of what behaviors coaches do actually engage in.

The *frequencies* and *percentages* provide a quantitative basis for describing the relative amounts of time spent by a coach in various types of behavior. It can be readily ascertained by looking at the percent column in Table 3 of the preceding chapter, that Coach A spent approximately 41 percent of his time in monitoring, 34 percent in instruction, 12 percent in praise and reward, and 4 percent in management behaviors.

These figures do not of course give a total picture of the coaching session, rather they must be interpreted in light of specific situational variables operating at the time. For example, instructional behaviors might be expected to occupy a more important position early in the season. It therefore is important to maintain a recording and coding log to assist with this interpretation. Appendix B presents an example of a coding sheet and explanatory comments.

Although the results mentioned above are taken from the combined sample data for Coach A, there may be a good case for looking at the details of specific practice

sessions. For example, perusal of the results for practice session 4 presented in Table 1 shows a greater percentage of time spent in management behaviors than is indicated in the combined data of Table 3. The recording log and coding sheet explanatory comments indicate that, on that particular occasion, the absence of several players necessitated the involvement of the coach in the drill practices. Consequently a greater number of management behaviors were coded for that practice session at the expense of the other categories.

One point of interest that arose in the discussion in chapter two concerning the inappropriate use of classroom observation instruments in physical education and sport settings, is supported by the data obtained in this study. Both coaches exhibited on average, approximately 38 percent of their total behavior as monitoring, which provides some support for the prevalence of such behaviors in coaching, and for the development of specific instruments for observing behavior in sport settings.

The *behavior clusters* group certain categories of behavior together to provide additional information about coaching. From the combined sample data for Coach B (Table 7) it can be seen that the coaching behaviors are almost equally divided between indirect (55 percent) and direct (45 percent) coaching. Affective behaviors account for approximately 13 percent. It must be noted that these clusters are not all mutually exclusive, such as positive-negative and direct-indirect coaching behaviors. For the

remainder it is sufficient to note that a certain percentage of behaviors can be labelled as primary or affective coaching. When data from a wide range of coaching situations has been analyzed these clusters may develop more meaning in terms of describing behavioral variation.

The *coaching style indices* calculated by using various behavior clusters as numerator and denominator, are expressed typically as ratios. Although the emphasis in this study is on descriptive rather than evaluative analysis, a comparison of the ratios provided in Tables 3 and 7 shows the capacity of these indices to describe variations in behavior. For example, there are differences in the primary coaching/manage index, the indirect/direct index, and the positive/negative index which may reflect different coaching climates or styles. Again the referral to the specific conditions of the different coaching environments, is essential in the interpretation of these figures.

An interesting addition to the original list of indices presented in the CBOS is the Primary coaching/manage index, which resulted from the preliminary data recording and coding phases of the study. It has been stated previously that a central function of coaching is instruction, and it may be that this index will be useful to coaches and coach trainers in terms of providing information about time which may be spent unnecessarily in management at the expense of actual instruction and feedback behaviors.

Further investigation of the usefulness of the

suggested sub-categories of instruction may result in other additions to these indices. For example, as more data becomes available it may prove worthwhile to compare the less-direct feedback, requests and questioning, with the direct command and direction sub-categories. This may provide information to supplement the current indirect/direct index which has proven to be useful in previous observational research (Flanders, 1967; Smith, 1974b).

Behavioral matrices which focus upon the sequential aspects of coaching behavior may be particularly useful to those concerned with the modification of coaching behavior. Knowledge as to which particular behaviors precede and which typically follow other behaviors may provide a basis for initiating and monitoring such changes (Smith, Smoll, and Curtis, 1979). Tables 4 and 8 provide examples of behavioral matrices for each coach with instructions as to the calculation of preceding and following behaviors. Such matrices may provide interesting information concerning a number of aspects of coaching behavior sequences. For example, Table 4 shows that Coach A spent approximately 5 percent of his coaching behavior in corrective feedback. Of the 204 observed instances of corrective feedback across the four samples recorded, 76 were followed by instruction. A similar finding was noted in Tharp and Gallimore's (1976) study of basketball coach John Wooden. They modified the categories of their observation instrument to include this type of re-instruction which they noted to be a significant

feature of this "master" coach's coaching behavior.

The steady state cells provide information concerning those behaviors which extend over the five second observation intervals employed in this study. Monitoring again heads the list for Coach A, while for Coach B instruction, monitoring and management behaviors show high incidences of extended presence.

For matrices as well as the percentages and clusters discussed above, more specific information can be obtained by referring to the appropriate coding sheet for individual practice sessions. In this way it is possible to determine for example, just how long some of these steady state behaviors actually persisted.

The example of a segment of coding from sample 4 for Coach A presented in Appendix B shows that of the 180 instances of monitoring, extended monitoring of twenty five seconds occurred once, twenty seconds six times, fifteen seconds eleven times, and ten seconds twenty-three times. These figures are not entirely accurate, as an inspection of Appendix B indicates that some behaviors only overlap a portion of the next five second interval.

Combined Data

The percentages, clusters and indices presented in Table 9, and the matrix in Table 10, which are based on combined data for both coaches over all samples, are presented here only as a preliminary suggestion for the way in which

data may be combined, for example, for coaches of a certain team sport at a given level. The temptation to do this must, however, be tempered by the importance of specific situational factors which become lost when data is combined in this way. It may be that the patterns of coaching behavior, even within the one sport and level, are so idiosyncratic or situationally specific that this approach will provide, at best, superficial and general information about trends in the data. Hence, the presentation of data in Tables 9 and 10 should be viewed as suggestive only, and in no way does the data indicate what might be considered as representative behaviors of team sport coaches at university level competition.

The Utility of the CBOS Analyses

Because of the preliminary nature of this study and lack of additional data available on the use of the CBOS, there is no quantitative evidence of support for the utility of the various analyses suggested by Smith (1978).

Intuitively, each of the various measures employed in the CBOS data analysis provides potentially useful information. This usefulness is based on their capacity for describing as accurately as structured observation allows, and as consistently as the instrument and observer(s) permits, what actually does occur in terms of practice session coaching behaviors. When valid instruments have been developed and tested, and research findings replicated,

then these measures may be potentially useful in providing information for the training of coaches, and in facilitating the monitoring of planned changes in coaching behaviors (Smith, 1978; Smith, Smoll and Curtis, 1979).

The results of this study do not provide any basis for suggesting the deletion of particular segments of the CBOS analyses.

It would be inappropriate for example, to focus upon the categories of behavior which were infrequently observed, and to use this as an indication of limited usefulness. The results presented in Table 1 show only one instance of recorded negative demonstration for Coach A during that particular practice sessions. Similarly, Table 5 indicates that only one example of negative demonstration was observed for Coach B during that single practice session. The combined results in Table 9 show only 0.1 percent of the 11,694 observed behaviors, and yet this cannot be interpreted as an indication that this particular type of behavior is insignificant and that the category should be deleted from the analysis.

It must be emphasized that the frequency of observed behavior may not necessarily be related to its significance. In the examples mentioned above it may be the case that the few instances of negative demonstration observed were of particular significance for the coach and/or players. This study provides no evidence of this, other than the explanatory comments on the coding sheet which describe

situations where a major change in the coach's behavior occurred because of the introduction of a new activity or an interruption such as a player injury.

CHAPTER VI

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

The purposes of the study were to examine the validity of the CBOS categories, to begin the development of a data base of videotape recorded practice session coaching behaviors, and to investigate the usefulness of the CBOS descriptive analyses in terms of providing an accurate and reliable picture of those coaching behaviors.

The coaching behaviors of the coaches were recorded over a total of eight practice sessions. After appropriate training, these behaviors were coded by the investigator according to the CBOS categories and ground rules, and computer analyzed to provide descriptive information concerning category percentages and frequencies, behavior clusters, coaching style indices, and sequence matrices.

The validity of the CBOS categories was examined in terms of the mutual exclusiveness of the categories and related reliability measures, and the range of behaviors accounted for by the categories. In this respect it was proposed that three aspects of reliability should be measured in a study of this nature. In addition, modifications were suggested to the category definitions to account for the range of observed behaviors, and preliminary suggestions were made for the further subdivision of the Instruct category.

The descriptive analyses of the CBOS generated data

yielded a number of measures which were considered useful in describing various aspects of coaching behavior during practice sessions. Description rather than evaluation was emphasized in the interpretation of the results.

Conclusions

On the basis of the results of this preliminary study the following conclusions are made:

1. The CBOS can be used in an accurate and reliable manner to code observed practice session coaching behaviors, and to provide useful descriptive information about the coaching process.
2. The validity of the CBOS categories is further enhanced by the inclusion of the additions suggested for increasing the range of behaviors accounted for, and the preliminary suggestions made for the further subdivision of the Instruct category.
3. The high frequencies of monitoring and instruction reported in the results support the need for the development of observational instruments, such as the CBOS, specifically designed for the study of coaching behavior. These instruments should include specific categories for the coding of various forms of instructional behavior.
4. The results from the various approaches to reliability measurement adopted in this study

indicate that measuring observer accuracy and stability, and inter-observer agreement are necessary features to be included in observational studies of this type.

5. The results of the study support the utility of the observational method, as a complimentary research technique to the more traditional trait approach, in the study of the coaching process.

Recommendations

It is recommended that the CBOS categories may be more effectively used to provide descriptive information about coaching if the suggestions concerning additions to the range of behaviors accounted for, further subdivision of the Instruct category, the sequence for observer training, and the various reliability measures, are adopted.

In addition, it is recommended that the observational method be pursued as an approach to the study of coaching.

The CBOS shows potential for further use in observational research which focuses upon practice session coaching behavior. The recommendations which follow approximate a suggested general sequence for future research employing the CBOS.

Modifications of the CBOS

The suggestions made in this study for the subdivision of the Instruct category and the inclusion of the additional coaching style index, require further investigation.

There may be additional modifications which will increase the utility of the descriptive measures. For example, as a result of the suggested changes to the Instruct category, another coaching style index might be developed to describe differences in coach-centred (command), and player-centred (feedback, questions, requests) behaviors.

<u>Training Procedures</u>

Future investigation of the development of a standardized training manual, perhaps incorporating an audio-visual module (Smith, et. al.,1979) is required. Emphasis has been placed in this study on an investigation of various aspects of reliability measurement, but because of the single-observer nature of this study, this should be extended to include more observers and other statistical procedures (Hollenbeck, 1978).

<u>Data Collection</u>

There is an obvious need for further development of a data base which includes samples of observed coaching behavior from a large number of coaches of different sports, at different levels of competition. These samples should also be collected at different stages of the competitive season and at different times in the weekly preparation for competition.

To assist with the coding of behavioral *effects* as specified in the CBOS ground rules, at least two video-cameras should be used in conjunction with a split screen

generator so that the coach and the player(s) with whom the interaction occurs, can be monitored.

There is also a need to involve the *subject* in observational research, in providing relevant feedback as to the accuracy of coding (Harre and Secord, 1972).

Behavioral Coding

Investigation of the utility and procedures required for coding the *significance* of the effect of practice session coaching behaviors of interest, is desirable. This will provide for useful information, in addition to the descriptive data generated by the analyses employed in this study, and for a more accurate representation of observed behavior.

REFERENCES

Adams, R.
1972 Observational studies of teacher role. In: N. Flanders (Ed.), Classroom Behavior of Teachers. Hamburg, West Germany: UNESCO Institute of Education, 440-459.

Amidon, E. J. and Hough, J. B. (Eds.)
1967 Interaction Analysis: Theory, Research and Application. Reading, Mass.: Addison-Wesley.

Anderson, W. G.
1971 Descriptive analytic research on teaching. Quest. 15, 1-8.

Anderson, R. M.
1972 A comparison of Bale's and Flander's systems of interaction analysis as research tools in small group instruction. Unpublished Ph.D. dissertation, The University of Alberta.

Anderson, W. G.
1976 Formal analysis. Mimeographed manuscript, Department of Physical Education, Teacher's College, Columbia University, New York.

Bany, M. A. and Johnson, L. V.
1964 Classroom Group Behavior. Toronto: Collier-Macmillan.

Barrett, K. R.
1969 A procedure for systematically describing teacher-student behavior in primary education. Unpublished Ph.D. dissertation, University of Wisconsin.

Beale, J.
1975 A preliminary model for the examination of intention through content analysis. Communication Research. 2(1), 86-96.

Berelson, B.
1952 Content Analysis in Communication Research. Glencoe, Illinois: The Free Press of Glencoe.

Biddle, B. J.
1967 Methods and concepts in classroom research. Review of Educational Research. 37, 337-357.

Blumer, H.
1967 Sociological analysis and the "variable". In: J. G. Manis and B. N. Meltzer (Eds.). Symbolic Interaction: A Reader in Social Psychology. Boston: Allyn and Bacon, 84-94.

Bookhout, E. C.
1967 An observational study of teaching behavior in relation to the social-emotional climate of physical education classes. Research Quarterly. 33, 336-347.

Bowers, J. W.
1970 Content analysis. In: P. Emmert and W. Brooks (Eds.). Methods of Research in Communication. Boston: Houghton Miflin, 291-314.

Bruner, J.
1966 Toward a Theory of Instruction. Cambridge, Mass.: Belknap Press.

Cartwright, D.
1975 Trait and other sources of variance in the S-R inventory of anxiousness. Journal of Personality and Social Psychology. 32,

Chelladuria, P. and Haggerty, T.
1978 Cited in: P. Chelladuria and A. V. Carron. Leadership, CAPHER, University of Calgary Press.

Cohen, J.
1960 A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 20(1): 37-46.

Cronbach, L. J.
1970 Essentials of Psychological Testing. 3rd. Edition, New York: Harper and Row.

Cronbach, L. J. and Meehl, P. E.
1955 Construct validity in psychological tests. Psychological Bulletin. 52(4), 281-302.

Danielson, R. R.
1974 Leadership in coaching: Description and evaluation. Unpublished Ph.D. dissertation, The University of Alberta.

Devereux, E. G.
1976 Backyard versus Little League baseball: Some observations on the impoverishment of children's games in contemporary America. In: A. Yiannakis, T. D. McIntyre, M. J. Melnick, and D. P. Hart (Eds.), Sport Sociology: Contemporary Themes. Dubuque, Iowa: Kendall-Hunt, 96-102.

Division of Educational and Research Services
1971 Flanders Interaction Analysis, Test 13. Computer Program Documentation, The University of Alberta.

Dougherty, N. J.
1971 A plan for the analysis of teacher-pupil interaction in physical education classes. Quest. 15, 39-49.

Endler, N. S. and Hunt, J. M.
1966 Sources of behavioral variance as measured by the S-R inventory of anxiousness. Psychological Bulletin. 65, 336-346.

Endler, N. S. and Magnussen, D. (Eds.)
1976 Interactional Psychology and Personality. New York: Wiley.

Fisher, A. C.
1977 Sport personality assessment: Adversary proceedings. In: C. O. Dodson, V. L. Katch, and J. Shick. (Eds.) Research and Practice in Physical Education. Champaign, Illinois: Human Kinetics.

Fisher, A. C., Horsfall, J. S. and Morris, H. H.
1977 Sport personality assessment: A methodological re-examination. International Journal of Sport Psychology. 8, 92-102

Fishman, S. E.
1970 A procedure for recording augmented feedback in physical education classes. Unpublished Ph.D. dissertation, Teachers College, Columbia University.

Fishman, S. and Anderson, W. G.
1971 Developing a system for describing teaching. Quest. 15, 9-16.

Flanders, N.
1965 Teacher Influence, Pupil Attitudes and Achievement. Washington: U.S. Department of Health, Education and Welfare.

Flanders, N.A.
1967 The problems of observer training and reliability. In: E. J. Amidon and J. B. Hough (Eds.) Interaction Analysis: Theory, Research and Application. Reading, Mass.: Addison-Wesley, 158-166.

Flanders, N.
1970 Analyzing Teacher Behavior. Reading, Mass.: Addison-Wesley.

Fodor, J.
1979 Personal communication.

Gentile, A.M.A.
1972 A working model of skill acquisition with application to teaching. Quest. 17, 3-23.

Golding, S. L.
1975 Flies in the ointment: Methodological problems in the analysis of percentages of variance due to persons and situations. Psychological Bulletin. 82, 278-288.

Good, C. B. (Ed.)
1973 Dictionary of Education. Third Edition. New York: McGraw-Hill.

Gravelle, L.
1977 Master coach and swim team: an ethnographic account. Unpublished Ph.D. dissertation, University of Alberta, 1977.

Harre, R. and Secord, P. F.
1972 The Explanation of Social Behavior. Oxford: Blackwell.

Hawkins, R. P. and Dotson, V. A.
1975 Reliability scores that delude: an Alice in Wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. In: E. Ramp and G. Semb (Eds.). Behavior Analysis: Areas of Research and Application. New Jersey: Prentice-Hall, 359-376.

Hendry. L. B.
1969 A personality study of highly successful and "ideal" swimming coaches. Research Quarterly. 40, 299-305.

Hilgard, E. R.
1964 A perspective on the relationship between learning theory and educational practices. In: E. R. Hilgard (Ed.) Theories of Learning and Instruction. The Sixty-third Yearbook of the National Society for the Study of Education, Part 1. Chicago: NSSE, 402-415.

Hollenbeck, A. R.
1978 Problems of reliability in observational research. In: G. P. Sackett (Ed.) Observing Behavior. Volume 2: Data Collection and Analysis Methods. Baltimore: University Park Press, 79-98.

Hough, J. B.
1967 An observational system for the analysis of classroom instruction. In: E. J. Amidon and J. B. Hough

(Eds.) Interaction Analysis: Theory, Research and Application. Reading, Mass.: Addison-Wesley, 150-157.

Johnson, S. M. and Bolstad, O. D.
1973 Methodological issues in naturalistic observation: some problems and solutions for field research. In: L. A. Hamerlynck, L. C. Handy and E. J. Mast (Eds.) Behavioral Change. Champaign, Illinois: Research Press, 7-67.

Kazdin, A. E.
1977 Artifact, bias and complexity of measurement: the ABC's of reliability. Journal of Applied Behavior Analysis. 10(1), 141-150.

Kroll, W.
1970 Current strategies and problems in personality assessment of athletes. In: L. E. Smith (Ed.) Psychology of Motor Learning. Chicago: Athletic Institute, 349-367.

Kuhn, T.
1962 The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Martens, R.
1975 The paradigmatic crisis in American sport personology. Sportwissenschaft. 5, 9-24.

Martin, J.
1976 Developing category observation instruments for the analysis of classroom behavior. Journal of Classroom Interaction. 12, 5-16.

Medley, D. M. and Mitzel, H. E.
1959 Some behavioral correlates of teacher effectiveness. Journal of Educational Psychology. 50, 239-246.

Medley, D. M. and Mitzel, H. E.
1963 Measuring classroom behavior by systematic observation. In: N. L. Gage (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally, 217-328.

Millson, L. and Christie, J.
1979 Hockey dropouts lose to pressure. The Globe and Mail, March 20.

Morgan, W. P.
1978 Sport personology: The credulous-skeptical argument in perspective. In: A. A. Ismail (Ed.) Proceedings of the Third Symposium on Integrated Development: Psycho-social Behavior of Sport and

Play. Indianapolis: Indiana State Board of Health, 218-227.

Nygaard, G.
1975 Interaction analysis in physical education classes. Research Quarterly. 46(3), 351-357.

Ogilvie, B. and Tutko, T. A.
1968 Self image and measured personality of coaches. Paper presented at the Second International Congress of Sport Psychology. Washington, D.C. November.

Orlick, T. D.
1974 The sports environment. A capacity to enhance - a capacity to destroy. Paper presented at the Sixty Canadian Symposium on Psychomotor Learning and Sports Psychology. Halifax, Nova Scotia, November.

Orlick, T. and Botterill, C.
1975. Every Kid Can Win. Chicago: Nelson-Hall.

Quereshi, M. Y.
1974 Purposes and procedures of validation. The Journal of Psychology. 88 , 77-89.

Ritzer, G.
1975 Sociology: A Multiple Paradigm Science. Boston: Allyn and Bacon.

Robbins, S. G.
1973 The development of an instrument to analyze teacher behavior in elementary school physical education. Unpublished Ph.D. dissertation, The University of Alberta.

Rosenshine, B. and Furst, N.
1973 The use of direct observation to study teaching. In: R. M. W. Travers (Ed.) Second Handbook of Research on Teaching. Chicago: Rand McNally, 122-183.

Rushall, B. S.
1973 Two observation schedules for sporting and physical education environments. Unpublished paper, Dalhousie University.

Rushall, B. S.
1978 Environment specific behavior inventories: Developmental procedures. International Journal of Sport Psychology. 9(2), 97-110.

Schwartz, S.
1972 A learning based system to categorize teacher behavior. Quest. 17, 52-55.

Scott, W. A.
1955 Reliability of content analysis: the case of nominal scale coding. Public Opinion Quarterly. 19(3), 321-325.

Simon, A. and Boyer, E. G. (Eds.)
1967 Mirrors for Behavior: An Anthology of Classroom Observation Instruments. Vols. 1-6. Philadelphia: Research for Better Schools.

Simon, A. and Boyer, E. G. (Eds.)
1970 Mirrors For Behavior: An Anthology of Classroom Observation Instruments. Vols. 7-14. Philadelphia: Research for Better Schools.

Simon, A. and Boyer, E. G.
1974 Mirrors For Behavior 111: An Anthology of Classroom Observation Instruments. Philadelphia: Research for Better Schools.

Smith, M. F. R.
1973 A preliminary case for classifying sports environments by participant objectives. Paper presented at the British Columbia Conference on Health and Physical Activity, Vancouver, British Columbia, November.

Smith, M. F. R.
1974a Parent communication and training by the coach. Proceedings of the International Symposium on the Art and Science of Coaching. Volume 2. Toronto, 147-155.

Smith, M. F. R.
1974b A psychologically based approach to teaching swimming. Unpublished Ph.D. dissertation. The University of Alberta.

Smith, M. F. R.
1975a Sport and physical activity as possible factors in the child's psychological and social development. Paper presented at Intersport One, a conference sponsored by the Sports and Fitness Branch of the Department of Recreation, Parks and Wildlife, Government of Alberta, Red Deer, Alberta, April.

Smith, M. F. R.
1975b Physical education, the sport sciences and Whitehead's dilemma. Mouvement: Seventh Canadian

Symposium on Psychomotor Learning and Sport Psychology. October, 227-236.

Smith, M. F. R.
1978 An Observational System for the Analysis of Coaching Behavior. Research Project No. 125. Health and Welfare Canada, Fitness and Amateur Sport Branch, June.

Smith, R. E., Smoll, F. L. and Hunt, E.
1977 A system for the behavioral assessment of athletic coaches. Research Quarterly. 48(2), 401-407.

Smith, R. E. and Smoll, F. L.
1978 Sport and the child. Conceptual and research perspectives. In: F. L. Smoll and R. E. Smith (Eds.). Psychological Perspectives in Youth Sports. Washington: Halsted Press, 3-13.

Smith, R. E., Smoll, F. L. and Curtis, W.
1978 Coaching behaviors in Little League Baseball. In: F. L. Smoll and R. E. Smith (Eds.), Psychological Perspectives in Youth Sports. Washington: Halsted Press, 173-201.

Smith, R. E., Smoll, F. L., Hunt, E., Curtis, W. and Coppel, D.B.
1978 Psychology and the bad news bears. Paper presented at NASPE Conference, Talahassie, Florida, Spring.

Smith, R. E., Smoll, F. L. and Curtis, W.
1979 Coach effectiveness training: A cognitive-behavioral approach to enhancing relationship skills in youth sport coaches. Journal of Sport Psychology. 1(1), 59-75.

Tharp, R. G. and Gallimore, R.
1976 What a coach can teach a teacher. Psychology Today. 75-78, January.

Usher, P. M.
1977 A descriptive analysis of coaching behavior. Unpublished Ph.D. dissertation, The University of Alberta.

Wankel, L. M. and Pabich, P.
1977 Coaching behavior scale. Unpublished manuscript, McMaster University, Hamilton, Ontario.

Webster's Third New International Dictionary.
1966 Springfield, Massachusetts: Merriam.

Weick, K. E.
1968 Systematic observational methods. In: G. Lindzey and E. Aronson (Eds.) The Handbook of Social Psychology, Volume 2. Reading, Mass.: Addison-Wesley, 357-451.

Wilson, S., Buzzell, N. and Jensen, M.
1975 Observational research: A practical tool. Physical Educator. 32(2), 90-93.

Wood, L.
1979. Personal communication, May 10.

APPENDICES

APPENDIX A

CBOS CATEGORY DEFINITIONS AND GROUND RULES

The category definitions and ground rules which follow are taken directly from Smith (1978, pp. 2-7). The *category definitions* are arranged under two major headings: indirect, and direct coaching behaviors.

Indirect Behaviors.

1. Monitor. The coach silently observes individuals, small groups, or the entire group as they practice or play. He or she may talk to themselves or show incidental nonverbal behavior while monitoring. Unless these behaviors are seen to have real effects on participants they should be ignored and considered to be external signs of the covert thought processes going on in the coach's mind during monitoring.
2. Praise and reward. The coach verbally or non-verbally praises, offers compliments, encourage-ment, acceptance, or agreement, with participant behavior. Statements may refer to present, past, or future behavior or performance. Nonverbal praise such as nodding, pats on the back, hugging, clapping as in applause, and other recognized gestures of approval are included. Examples of verbal praise, with appropriate

tone of voice: "Great!" "Yes, good work!" "Keep going, you'll get it!" "OK! OK!"

Direct Behaviors.

3. Instruct. Statements about what to do, or how and why to do it. Questions to elicit recall of earlier instruction or to get participant ideas on performance technique, tactics, attitudes, or practice activity. This category includes requests and commands. Examples: "Hold the hand in this position." How could you try to counter a move like that by your opponent?" "Be alert all the time." "OK, go ahead." "Stop. Run in and listen to this." Do you remember what we said about pacing yourself?"
4. Demonstration, positive. The coach, or a participant at the request of the coach, demonstrates how an action is to be carried out. See the ground rules following for an explanation of how to record combined verbal instruction and demonstration.
5. Demonstration, negative. The coach, or a participant at the request of the coach, demonstrates an error in action or an action to be avoided. Again, see the ground rules for recording mixed instruction and negative demonstration.

6. <u>Corrective feedback</u>. Statements or guidelines by the coach that follow direct observation of performance and which are intended to point out errors or inappropriate aspects of performance. Such statements are essentially information, as contrasted with praise or scolds which tend to convey an emotional reaction in the sense of positive valuation or negative valuation of the coach to an observed performance. Corrective feedback is restricted to knowledge or skill areas in which performance can be considered correct or appropriate by definition, or as a matter of fact, by convention. Examples: "Your head is too far forward." "Can you get more weight onto your left foot?" "That is not how we play a two-on-one situation" (stated in a matter of fact tone of voice). "You are unstable because your feet are too close together, spread them, widen your base."

7. <u>Hustles</u>. Statements or actions by the coach to activate or intensify previously instructed behaviors. Among the actions that can be used as hustles are vigorous, short, rapid clapping (not intended as applause); a quick jabbing motion of the forefinger directed at an individual or group; a quick lunging motion

of the shoulders or body; and any other nonverbal behavior that is recognized by participants as calling for more effort or intensity. Examples of verbal hustles: "Go! Go!" "Drive harder!" "Move it!" "C'mon!" "Hustle!"

8. Scold. Statements intended to scold, criticize, or reject the performance, behavior, feelings, or ideas of participants. This category includes sarcastic statements and expressions of displeasure, both verbal and nonverbal. Among nonverbal scolds are stamping the feet in disgust; a rapid whirl or pivot away from the offending action; a quick movement of the hands to the hips; slapping the forehead as the head is thrown back; a threatened or actual kick, punch, choking, or slashing action; a so-called raspberry or other unpleasant sound, such as a prolonged hiss; and any other action recognized by participants as intended to scold, criticize or reject. Examples of verbal scolds (with an appropriate tone of voice): "Terrible!" "Stop that right now!" "Get out!" "Lovely, just bloody lovely!" "What the hell do you think you're doing?" "How can you be so stupid?" (lazy, selfish, cowardly, etc.).

9. Management. The coach engages in verbal or nonverbal activity associated with management, housekeeping, routine procedures, announcements, arrangements, other than instruction or arranging or grouping participants for instruction or practice activity. Among management behaviors are taking attendance; making arrangements with grounds keepers, janitors, or others working or playing in the area; collecting money from participants; giving details related to schedules or travel; arranging car pools; distributing or collecting game uniforms; and so on.

10. Other. Any coaching behaviors displayed which do not fit into any of the other nine categories. Statements that cannot be understood or occasions where the noise level is so high that the coach cannot be heard over it are recorded here. Chatting or visiting with others, or attending to personal matters unrelated to the coaching or management tasks at hand, fall in this category.

Ground rules for using the system. As with many observation systems it is necessary to outline a set of ground rules to set standard conditions for observing behaviors

to provide some formal rules to use in assigning ambiguous behaviors to the most appropriate category.

Rule 1. Numbers corresponding to the ten categories are recorded by an observer every five seconds. When more than one category of behavior is observed during a five second interval all categories observed are recorded. The numbers indicating categories of observed behavior are recorded in vertical columns of twelve digits each column, thus representing one minute of observation. If any one category of behavior is observed to continue then the number representing the category is repeated every five seconds. Each practice or workout or portion of one that is observed begins and ends by arbitrary convention by recording a 10.

Rule 2. Because the system is descriptive, not evaluative, the observer should not try to guess the intent of the coach, but rather to assign behaviors to categories according to how he or she perceives their *effect on participants*. For example, a comment may sound like a hustle, but the participants clearly react with annoyance and embarrassment which would indicate it had been received as a scold. Then

it should be recorded as a scold.

Rule 3. Tone of voice and other aspects of nonverbal behavior must be used in assigning behaviors to categories. The same question, for example, asked in different tones of voice, with appropriate gestures or facial expressions might be properly categorized as instruction (as defined above), as a praise, as a scold, or even as corrective feedback.

Rule 4. Demonstrations, either positive or negative, are usually accompanied by verbal instruction. When a demonstration is carried on silently and not described as the action is carried out a four (4) is recorded if positive, a five (5) if negative, at each five second interval as long as the demonstration continues. A demonstration may shift back and forth from positive to negative, which would be indicated by recording 4's and 5's as appropriate. If the demonstration is accompanied simultaneously by verbal instruction, record a 3 *and* either a 4 or 5 for *each* five second interval during which this is continued.

Rule 5. Laughter by the coach in response to a participant is praise if it shares positive feelings or reassures, but is a scold if it

rejects or ridicules.

Rule 6. A phrase such as "OK" or "alright" may be a praise or scold depending on the context or when used in a matter-of-fact way as in, "Well that's done, let's move on to the next thing," it would be an instruction.

Rule 7. Saying a participant's name may be merely part of an instruction (when said matter-of-factly), a praise, a scold, a hustle, or even corrective feedback, depending on the tone of voice and other nonverbal behavior the coach displays when he or she says it.

Rule 8. In environments when participants are spread out over large areas, or where vision or hearing are difficult (as in a pool), normal intervals of time required in waiting because of such problems should be recorded as monitoring, if the coach is silent. However, during the period the coach waits to get the attention of the group, or as they assemble, he or she may produce behaviors clearly in any of the other nine categories.

APPENDIX B

TABLE II

CALCULATION OF CBOS DESCRIPTIVE MEASURES

Sample Coding Sheet

5 sec.	Each column approximately one minute										Sport.............
interval	1	2	3	4	5	6	7	8	9	10	Date..............
1	10	9	9	9	1/3	1	1	3/1	1/2	3/1	Start....Finish.....
2	3	9	2/9	9	1	1/2	2/3	1	3/9	3/1	Season Begins.......
3	3	2/9	9	2/9	9	1	3	1/2	9	3/1	Season Ends.........
4	3	9/2	9	9	1/2	1	3	1/2	3/1/2	3/1	
5	9	9	9	9	1/7	1/3	3	3/1	3/1	1/2/3	Tape ooo
6	9	9	9/2	9	1	1	9	6	3/1	3/1	"Yeh, I'll have to
7	9	9/2	3	9	1	1	9	3/1	3/1	3/1	run..."
8	9	9	9	3/1	1/2	1	3/1	3/1	3/1	1	Ends...
9	3/9	2/9	9	1/3	1	1	3/1	3/1	3/1	1	"Back out, Green
10	9	9	2/9	1	3	2/3	3/1	3/1/6	1	1	ball..."
11	9	9	9	1/2/3	1	1	3/1	6	1	1/3	LENGTH 10.0 min
12	9	9	9	1	1	2/3/7	3/1	3/1	1	10	RATE/MIN. 17.5

EXPLANATORY COMMENTS: (1) Coach relieved from participation by arrival of additional player (2) End lay up drill.

APPENDIX B cont'd

TABLE 12

CALCULATION OF CBOS DESCRIPTIVE MEASURES

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	61	34.9
2. Praise	20	11.4
3. Instruct	44	25.1
4. Demo +	0	0
5. Demo -	0	0
6. Correct	3	1.7
7. Hustle	2	1.1
8 Scold	0	0
9. Manage	43	24.6
10. Other	2	1.1
TOTALS	175	100

Total practice time - 10.0 minutes
Observations/minute - 17.5

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (47/175)
   100 - 26.9
2. Affective Coaching (A)
   categories 2, 7, 8 (22/175)
   100 = 12.6
3. Indirect Coaching (I)
   categories 1, 2 (81/175)
   100 - 46.3
4. Direct Coaching (D)
   categories 3-9 (81/175)
   100 = 52.6
5. Positive Coaching (P)
   categories 2, 4 (20/175)
   100 = 11.4
6. Negative Coaching (N)
   categories 5, 8 (0/175)
   100 = 0

Coaching Style Indices

1. Primary coaching/affective
   C/A = 47/22 = 2.14
2. Primary coaching/monitor
   C/M = 47/61 = 0.77
3. Primary coaching/manage
   C/Ma = 47/43 = 1.09
4. Indirect/direct
   I/D = 81/92 = 0.88
5. Positive/negative
   P/N = 20/0 = 20.0

APPENDIX B cont'd

TABLE 13

CALCULATION OF CBOS BEHAVIORAL MATRIX

THE SECOND NUMBER IN THE PAIR LOCATES THE COLUMN

THE FIRST NUMBER IN THE PAIR LOCATES THE ROW

	MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor 1											
Praise 2											
Instruct 3			✓✓						✓✓		
Demo + 4											
Demo - 5											
Correct 6											
Hustle 7											
Scold 8											
Manage 9			✓						✓✓✓✓✓✓✓		
Other 10			✓								
TOTAL											
%											

Behavioral Matrix

Refer to first column of Table 11

1st pair		3rd pair		5th pair		7th pair		9th pair		11th pair		
10	3	3	3	9	9	9	9	3	9	9	9	9
	2nd pair		4th pair		6th pair		8th pair		10th pair		12th pair	

Each becomes a single tally in the appropriate cell of the matrix:

1st pair - row 10, column 3
2nd pair - row 3, column 3
3rd pair - row 3, column 3

NOTE: The last tally in column one of Table 11 (a9), and the first tally in column 2 (a9) become the 13th pair.

APPENDIX C

VIDEOTAPING HINTS

The first ten points which follow are taken from Smith (1978, pp. 62-63). The rest result from the investigator's experience during this study.

1. Meet with coach day or two before taping, or drop in to watch workout so you see the general use of space, orientation within space, and get a feel for movement and interaction patterns of greatest significance.

Ask coach if anything unusual can be expected in session you will tape. e.g. a special drill, anything extremely different from usual.

2. Emphasize coach should work as usual. Do not inhibit movement, actions, talk because of taping. Some reasonable effort to face general direction of camera most of the time is all that is expected. If location of camera (it may have to be moved) is discussed prior to workout, this should not be a problem. The coach should not hesitate to turn his back on the camera at any time he wants to. No special effort should be necessary with respect to voice. The coach should use his regular language and volume. The equipment is flexible enough to record what is said.

3. When one camera is in use the best orientation is to show front or profile of coach most or all the time, and show front or profile of fair numbers of athletes as much

of time as possible. When a choice must be made -- favor a better view of the coach.

4. If any natural back light is available, try to arrange with coach to orient things so that this light is behind the camera to best degree possible without compromising normal practice activity.

5. If movements of athletes block out view of coach with any regularity, elevate camera on tripod and/or onto a bench, table, or platform, so that view is reasonably unobstructed.

6. Have coach (not cameraman) explain to athletes briefly prior to workout what is going on. If coach or athletes want to ask him questions, they should feel free to do so. This will lower the level of excitement and restore things to close to normal.

If athletes look at or talk about the camera to any significant degree the coach should distract them or try subtle means to direct their attention to the practice task at hand. If that doesn't work then he should use more direct means to re-focus them on the practice. Coach should avoid looking at or making remarks about the camera.

7. If coach uses extensive or significant amounts of nonverbal to control, direct, or provide feedback, try to get representative samples of this in medium close-up, including several athletes if possible. If not, then tape the coach only.

During extended practice activity if there is

neither verbal nor nonverbal behavior of any significance, stop the tape until something starts, unless it has been decided to tape the whole practice or specified time segments.

8. Talk to the coach about how the workout ends. If he calls them together for a talk, summary, or any special instructions (except merely dismissing them), save 5-10 mins. of tape to get this.

9. If possible arrange for additional floodlights for areas with poor lighting. If in doubt, get sample taping to check picture quality before data collection begins.

10. Take about 15 secs. picture to title each tape: date, hour, place, coaches' names, at beginning of each tape.

11. Be very familar with use of the videotaping equipment. A mistake in setting the VTR switch to "camera" instead of "line" may result in an hour of tape with sound but no picture.

12. Check on maintenance of equipment, particularly the cleaning of the VTR heads and the condition of all leads. Remember, a videotape recorded on a VTR with dirty heads is useless for observational analysis. Included here is the charge condition of the batteries if a remote microphone is used.

13. Preview the tapes as soon after recording as possible. In this way trouble spots can be located before too much time and effort is wasted.

14. During previewing, tapes should be labelled with a pre-arranged indentification code, and some comments recorded in a log describing the general quality of the recording. These comments should supplement the general descriptive comments recorded in the log before and during filming which should describe the conditions, general outline, and significant features of the practice session. These records are valuable during subsequent coding. Part of the previewing should also include establishing a 000 starting point for replay. To assist here, the previewer should roll a consistent amount of tape onto the take-up reel and, as soon as the picture becomes stable look for some significant feature to pinpoint a reliable 000 start on the tape counter. This is important for training and reliability stages of an observational study.

15. The person filming practice session coaching behavior should give consideration to the protection of video equipment from moving players or balls. A moveable blackboard with a reasonable base of support may be useful.

16. Under ideal conditions where appropriate equipment is available, a cordless microphone should be used so that the coach's movement is less inhibited. A small clip-on microphone attached to a small pocket-size transmitter provides for free movement and an excellent reception range. Split screen techniques, involving a second camera focused upon the participant(s) and a special effects generator which connects the two cameras to a single VTR, allow for

the two camera images to be recorded and thus a wide range of combinations of behavior.

APPENDIX D

INITIAL CONTENT ANALYSIS CATEGORIES

The following list of categories was submitted to several experienced teacher/coaches for comment. The instructions were as follows:

The following is a list of some suggested sub-categories of the process of instruction. In the light of your experience in both teaching and coaching, can you suggest any modifications to this list?

1. COMMAND, ORDER, DIRECT.

Includes directions, orders and commands to which compliance is expected. Examples: "Hold your hand in this position." "Offense to that end, defense stay here." This category may also include nonverbal actions by the coach indicating to the players a direction to run, or when to start a drill (such as when the coach begins a drill by throwing or serving the ball).

2. LECTURE.

Includes giving facts or opinions about content or procedure. Examples: "In this drill it is important that this player runs right through to the baseline." "The key to our defense is the blocking power of our three front line players."

3. REQUEST.

Includes "toned-down" (compared to commands) statements which ask the player to do something. These

APPENDIX D cont'd

statements are often made quietly, person to person. Examples: "Now I would like you to try to..." After a player answers a previous question - "Well let's try this..."

4. EXPLAIN, EXPAND.

Includes statements which clarify, develop or expand a previously stated idea or suggestion provided by the coach. These statements may often follow a previous command or request. Example: "Running this offense properly means that our back line players will need to come in behind the spikers..."

5. QUESTION, IMPLIED QUESTION.

Includes questions to which answers are expected, but do not serve the function of other categories. These questions usually concern content or procedure. Example: "How could you counter a move like that?" May also include an extended pause with raised eyebrows (implied question).

6. RESPONSE TO PLAYER QUESTION.

Includes direct answers to player questions. Such answers may give information or opinion, but must be responses that answer or are directed toward answering player questions. Example: "No, you need to remain closer to the baseline."

APPENDIX E

TABLE 14

Coach A, Sample 1

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	453	45.7
2. Praise	118	11.9
3. Instruct	333	33.6
4. Demo +	2	0.2
5. Demo -	1	0.1
6. Correct	49	4.9
7. Hustle	21	2.1
8. Scold	3	0.3
9. Manage	9	0.9
10. Other	2	0.2
TOTALS	991	100

Total practice time - 57.0 minutes
Observations/minute - 17.4

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (385/991)
   100 = 38.9%
2. Affective Coaching (A)
   categories 2, 7, 8 (142/991)
   100 = 14.3%
3. Indirect Coaching (I)
   categories 1, 2 (571/991)
   100 = 57.6%
4. Direct Coaching (D)
   categories 3-9 (418/991)
   100 = 42.2%
5. Positive Coaching (P)
   categories 2, 4 (120/991)
   100 = 12.1%
6. Negative Coaching (N)
   categories 5, 8 (4/991)
   100 = 0.4%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 385/142 = 2.71
2. Primary coaching/monitor
   C/M = 385/453 = 0.85
3. Primary coaching/manage
   C/Ma = 385/9 = 42.77
4. Indirect/direct
   I/D = 571/418 = 1.37
5. Positive/negative
   P/N = 120/4 = 30.0

APPENDIX E. cont'd

Coach A, Sample 1

BEHAVIORAL MATRIX.

TABLE 15

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	209	64	147	0	0	21	6	2	3	1	453
Praise	2	37	13	56	0	0	8	2	1	1	0	118
Instruct	3	167	35	106	2	0	11	11	0	1	0	333
Demo +	4	0	0	0	0	1	1	0	0	0	0	2
Demo -	5	1	0	0	0	0	0	0	0	0	0	1
Correct	6	19	3	18	0	0	7	1	0	1	0	49
Hustle	7	16	0	3	0	0	0	1	0	1	0	21
Scold	8	1	1	1	0	0	0	0	0	0	0	3
Manage	9	3	1	1	0	0	1	0	0	2	1	9
Other	10	0	1	1	0	0	0	0	0	0	0	2
TOTAL		453	118	333	2	1	49	21	3	9	2	991
%		45.7	11.9	33.6	0.2	0.1	4.9	2.1	0.3	0.9	0.2	100

APPENDIX E cont'd

TABLE 16

Coach A, Sample 2

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	442	36.6
2. Praise	158	13.1
3. Instruct	456	37.7
4. Demo +	18	1.5
5. Demo -	0	0.0
6. Correct	67	5.5
7. Hustle	37	3.1
8. Scold	2	0.2
9. Manage	27	2.2
10. Other	2	0.2
TOTALS	1209	100

Total practice time 60.0 minutes
Observations/minute 20.15

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (541/1209)
   100 - 44.8%
2. Affective Coaching (A)
   categories 2, 7, 8 (197/1209)
   100 = 16.3%
3. Indirect Coaching (I)
   categories 1, 2 (600/1209)
   100 = 49.6%
4. Direct Coaching (D)
   categories 3-9 (607/1209)
   100 - 50.2%
5. Positive Coaching (P)
   categories 2, 4 (176/1209)
   100 = 14.6%
6. Negative Coaching (N)
   categories 5, 8 (2/1209)
   100 - 0.2%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 541/197 = 2.75
2. Primary coaching/monitor
   C/M = 541/442 = 1.22
3. Primary coaching/manage
   C/Ma = 541/27 = 20.04
4. Indirect/direct
   I/D = 600/607 = 0.99
5. Positive/negative
   P/N = 176/2 = 88.0

APPENDIX E cont'd

TABLE 17

Coach A , Sample 2
BEHAVIORAL MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	133	82	183	0	0	26	15	1	2	0	442
Praise	2	58	11	75	0	0	5	3	0	5	1	158
Instruct	3	202	56	141	12	0	19	18	0	8	0	456
Demo +	4	2	0	13	3	0	0	0	0	0	0	18
Demo -	5	0	0	0	0	0	0	0	0	0	0	0
Correct	6	21	7	21	3	0	14	0	0	1	0	67
Hustle	7	16	2	14	0	0	2	1	1	0	1	37
Scold	8	0	0	1	0	0	1	0	0	0	0	2
Manage	9	9	0	7	0	0	0	0	0	11	0	27
Other	10	1	0	1	0	0	0	0	0	0	0	2
TOTAL		442	158	456	18	0	67	37	2	27	2	1209
%		36.6	13.1	37.7	1.5	0.0	5.5	3.1	0.2	2.2	0.2	100

APPENDIX E cont'd

TABLE 18

Coach A, Sample 3.

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	467	44.3
2. Praise	120	11.4
3. Instruct	381	36.1
4. Demo +	2	0.2
5. Demo -	0	0.0
6. Correct	39	3.7
7. Hustle	29	2.8
8. Scold	3	0.3
9. Manage	11	1.0
10. Other	2	0.2
TOTALS	1054	100

Total practice time - 58.0 minutes
Observations/minute - 18.2

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (422/1054)
   100 = 40.0%

2. Affective Coaching (A)
   categories 2, 7, 8 (152/1054)
   100 = 14.4%

3. Indirect Coaching (I)
   categories 1, 2 (587/1054)
   100 = 55.7%

4. Direct Coaching (D)
   categories 3-9 (465/1054)
   100 = 44.1%

5. Positive Coaching (P)
   categories 2, 4 (122/1054)
   100 = 11.6%

6. Negative Coaching (N)
   categories 5, 8 (3/1054)
   100 = 0.3%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 422/152 = 2.78

2. Primary coaching/monitor
   C/M = 422/467 = 0.90

3. Primary coaching/manage
   C/Ma = 422/11 = 38.36

4. Indirect/direct
   I/D = 587/465 = 1.26

5. Postive/negative
   P/N = 122/3 = 40.67

APPENDIX E cont'd

TABLE 19

Coach A, Sample 3

BEHAVIORAL MATRIX

	MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor 1	192	68	185	0	0	12	9	1	0	0	467
Praise 2	31	9	67	0	0	6	4	1	2	0	120
Instruct 3	222	41	91	0	0	8	13	1	4	1	381
Demo + 4	0	0	1	0	0	1	0	0	0	0	2
Demo - 5	0	0	0	0	0	0	0	0	0	0	0
Correct 6	8	0	20	1	0	9	1	0	0	0	39
Hustle 7	12	2	12	0	0	1	2	0	0	0	29
Scold 8	0	0	2	0	0	1	0	0	0	0	3
Manage 9	2	0	3	0	0	0	0	0	5	1	11
Other 10	0	0	0	1	0	1	0	0	0	0	2
TOTAL	467	120	381	2	0	39	29	3	11	2	1054
%	44.3	11.4	36.1	0.2	0.0	3.7	2.8	0.3	1.0	0.2	100

APPENDIX E cont'd

TABLE 20

Coach A, Sample 4

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	415	40.0
2. Praise	108	10.4
3. Instruct	319	30.8
4. Demo +	4	0.4
5. Demo -	1	0.1
6. Correct	49	4.7
7. Hustle	16	1.5
8. Scold	12	1.2
9. Manage	111	10.7
10. Other	2	0.2
TOTALS	1037	100

Total practice time - 60.0 minutes
Observations/minute - 17.3

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (373/1037)
   100 = 40.0%
2. Affective Coaching (A)
   categories 2, 7, 8 (136/1037)
   100 = 13.1%
3. Indirect Coaching (I)
   categories 1, 2 (523/1037)
   100 = 50.4%
4. Direct Coaching (D)
   categories 3-9 (512/1037)
   100 = 49.4%
5. Positive Coaching (P)
   categories 2, 4 (112/1037)
   100 = 10.8%
6. Negative Coaching (N)
   categories 5, 8 (13/1037)
   100 = 1.3%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 373/136 = 2.74
2. Primary coaching/monitor
   C/M = 373/415 = 0.90
3. Primary coaching/manage
   C/Ma = 373/111 = 3.36
4. Indirect/direct
   I/D = 523/512 = 1.02
5. Positive/negative
   P/N = 112/13 = 8.62

APPENDIX E cont'd

TABLE 21

Coach A, Sample 4

BEHAVIORAL MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	158	71	150	0	0	16	5	4	11	0	415
Praise	2	33	7	46	0	0	8	1	0	13	0	108
Instruct	3	175	19	75	3	0	8	9	1	28	1	319
Demo +	4	0	0	3	0	0	1	0	0	0	0	4
Demo -	5	0	0	0	0	0	1	0	0	0	0	1
Correct	6	12	2	17	1	1	1	0	2	3	0	49
Hustle	7	7	0	6	0	0	1	1	0	1	0	16
Scold	8	2	0	2	0	0	2	0	5	1	0	12
Manage	9	27	9	19	0	0	1	0	0	54	1	111
Other	10	1	0	1	0	0	0	0	0	0	0	2
TOTAL		415	108	319	4	1	49	16	12	111	2	1037
%		40.0	10.4	30.8	0.4	0.1	4.7	1.5	1.2	10.7	0.2	100

APPENDIX E cont'd

TABLE 22

Coach B, Sample 1

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	736	37.9
2. Praise	129	6.6
3. Instruct	818	42.1
4. Demo +	3	0.2
5. Demo -	2	0.1
6. Correct	16	0.8
7. Hustle	89	4.6
8. Scold	9	0.5
9. Manage	111	5.7
10. Other	31	1.6
TOTALS	1944	100

Total practice time - 90.0 minutes
Observations/minute - 21.6

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (839/1944)
   100 = 43.2%
2. Affective Coaching (A)
   categories 2, 7, 8 (227/1944)
   100 = 11.7%
3. Indirect Coaching (I)
   categories 1, 2 (865/1944)
   100 = 44.5%
4. Direct Coaching (D)
   categories 3-9 (1048/1944)
   100 = 53.9%
5. Positive Coaching (P)
   categories 2, 4 (132/1944)
   100 = 6.8%
6. Negative Coaching (N)
   categories 5, 8 (11/1944)
   100 = 0.6%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 839/227 = 3.70
2. Primary coaching/monitor
   C/M = 839/736 = 1.14
3. Primary coaching/manage
   C/Ma = 839/111 = 7.56
4. Indirect/direct
   I/D = 865/1048 = 0.83
5. Positive/negative
   P/N = 132/11 = 12.00

APPENDIX E cont'd

TABLE 23

Coach B, Sample 1

BEHAVIORAL MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	127	72	471	1	1	10	28	6	12	8	736
Praise	2	18	4	98	0	0	0	7	0	1	0	128
Instruct	3	540	40	168	0	0	4	48	3	14	1	818
Demo +	4	0	0	3	0	0	0	0	0	0	0	3
Demo -	5	0	0	0	2	0	0	0	0	0	0	2
Correct	6	1	0	13	0	1	0	1	0	0	0	16
Hustle	7	31	13	39	0	0	2	3	0	1	0	89
Scold	8	1	0	6	0	0	0	2	0	0	0	9
Manage	9	13	0	16	0	0	0	0	0	12	12	111
Other	10	5	0	4	0	0	0	0	0	10	10	31
TOTAL		736	129	818	3	2	16	89	9	111	31	1944
%		37.9	6.6	42.1	0.2	0.1	0.8	4.6	0.5	5.7	1.6	100

APPENDIX E cont'd

TABLE 24

Coach B. Sample 2.

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	658	32.5
2. Praise	198	9.8
3. Instruct	793	39.2
4. Demo +	0	0.0
5. Demo -	0	0.0
6. Correct	40	2.0
7. Hustle	161	8.0
8. Scold	25	1.2
9. Manage	143	7.1
10. Other	5	0.2
TOTALS	2023	100

Total practice time - 94.0 minutes
Observations/minute - 21.4

Behavior Cluster Measures

1. Primary Coaching (P)
   categories 3-6 (833/2023)
   100 = 41.2%
2. Affective Coaching (A)
   categories, 2, 7, 8 (384/2023)
   100 = 19.0%
3. Indirect Coaching (I)
   categories 1, 2 (856/2023)
   100 = 42.3%
4. Direct Coaching (D)
   categories 3-9 (1162/2023)
   100 = 57.4%
5. Positive Coaching (P)
   categories 2, 4 (198/2023)
   100 = 9.8%
6. Negative Coaching (N)
   categories 5, 8 (25/2023)
   100 = 1.2%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 833/384 = 2.17
2. Primary coaching/monitor
   C/M = 833/658 = 1.27
3. Primary coaching/manage
   C/Ma = 833/143 = 5.83
4. Indirect/direct
   I/D = 856/1162 = 0.74
5. Positive/negative
   P/N = 198/25 = 7.92

APPENDIX E cont'd

TABLE 25

Coach B, Sample 2

BEHAVIORAL MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	137	87	328	0	0	19	69	7	10	1	658
Praise	2	45	7	127	0	0	2	12	0	5	0	198
Instruct	3	410	71	217	0	0	9	69	7	9	1	793
Demo +	4	0	0	0	0	0	0	0	0	0	0	0
Demo -	5	0	0	0	0	0	0	0	0	0	0	0
Correct	6	5	1	27	0	0	4	1	1	0	1	40
Hustle	7	45	25	69	0	0	6	7	7	0	1	161
Scold	8	4	2	14	0	0	0	3	3	0	0	25
Manage	9	10	5	7	0	0	0	0	0	119	1	143
Other	10	2	0	4	0	0	0	0	0	0	0	5
TOTAL		658	198	793	0	0	40	161	25	143	5	2023
%		32.5	9.8	39.2	0.0	0.0	2.0	8.0	1.2	7.1	0.2	100

APPENDIX E cont'd

TABLE 26

Coach B, Sample 3

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	755	36.5
2. Praise	189	9.1
3. Instruct	831	40.1
4. Demo +	1	0.0
5. Demo -	1	0.0
6. Correct	32	1.5
7. Hustle	28	1.4
8. Scold	5	0.2
9. Manage	224	10.8
10. Other	4	0.2
TOTALS	2070	100

Total practice time - 105.0 minutes
Observations/minute - 19.7

Behavior Cluster Measures

1. Primary Coaching (C)
   categories 3-6 (865/2070)
   100 = 41.8%

2. Affective Coaching (A)
   categories 2, 7, 8 (222/2070)
   100 = 10.73%

3. Indirect Coaching (I)
   categories 1, 2 (944/2070)
   100 = 45.6%

4. Direct Coaching (D)
   categories 3-9 (190/207)
   100 = 54.2%

5. Positive Coaching (P)
   categories 2, 4 (190/2070)
   100 = 0.9%

6. Negative Coaching (N)
   categories 5, 8 (6/2070)
   100 = 0.3%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 865/222 = 3.90

2. Primary coaching/monitor
   C/M = 865/755 = 1.15

3. Primary coaching/manage
   C/Ma = 865/224 = 3.86

4. Indirect/direct
   I/D = 944/1122 = 0.84

5. Positive/negative
   P/N = 190/6 = 31.67

APPENDIX E cont'd

TABLE 27

Coach B, Sample 3.

BEHAVIORAL MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	99	172	421	0	0	24	18	4	15	2	755
Praise	2	13	4	166	0	0	2	0	1	2	1	189
Instruct	3	629	9	172	1	0	2	7	0	11	0	831
Demo +	4	0	0	0	0	1	0	0	0	0	0	1
Demo -	5	0	0	1	0	0	0	0	0	0	0	1
Correct	6	1	0	27	0	0	4	0	0	0	0	32
Hustle	7	1	3	21	0	0	0	2	0	1	0	28
Scold	8	0	0	4	0	0	0	1	0	0	0	5
Manage	9	12	1	16	0	0	0	0	0	194	1	224
Other	10	0	0	3	0	0	0	0	0	1	0	4
TOTAL		755	189	831	1	1	32	28	5	224	4	2070
%		36.5	9.1	40.1	0.0	0.0	1.5	1.4	0.2	10.8	0.2	100

APPENDIX E cont'd

TABLE 28

Coach B, Sample 4.

CBOS Category Percentages

CATEGORY	COUNT	PERCENT
1. Monitor	555	40.6
2. Priase	87	6.4
3. Instruct	493	36.1
4. Demo +	1	0.1
5. Demo -	1	0.1
6. Correct	28	2.0
7. Hustle	9	0.7
8. Scold	2	0.1
9. Manage	180	13.2
10. Other	10	0.7
TOTALS	1366	100

Total practice time - 68.0 minutes
Observations/minute - 20.1

Behavior Cluster Measures

1. Primary Coaching (P)
   categories 3-6 (523/1366)
   100 = 38.3%
2. Affective Coaching (A)
   categories 2, 7, 8 (98/1366)
   100 = 7.2%
3. Indirect Coaching (I)
   categories 1, 2 (642/1366)
   100 = 47.0%
4. Direct Coaching (D)
   categories 3-9 (714/1366)
   100 = 52.3%
5. Positive Coaching (P)
   categories 2, 4 (88/1366)
   100 = 6.4%
6. Negative Coaching (N)
   categories 5, 8 (3/1366)
   100 = 0.2%

Coaching Style Indices

1. Primary coaching/affective
   C/A = 523/98 = 5.34
2. Primary coaching/monitor
   M/A = 523/555 = 0.94
3. Primary coaching/manage
   C/Ma = 523/180 = 2.91
4. Indirect/direct
   I/D = 642/714 = 0.90
5. Positive/negative
   P/N = 88/3 = 29.33

APPENDIX E cont'd

TABLE 29

Coach B, Sample 4

BEHAVIOR MATRIX

		MONITOR 1	PRIASE 2	INSTRUCT 3	DEMO + 4	DEMO - 5	CORRECT 6	HUSTLE 7	SCOLD 8	MANAGE 9	OTHER 10	TOTAL
Monitor	1	161	60	306	0	0	16	4	1	4	3	555
Praise	2	25	1	60	0	0	0	0	0	0	1	87
Instruct	3	353	25	95	0	0	7	4	1	7	1	493
Demo +	4	0	0	1	0	0	0	0	0	0	0	1
Demo -	5	0	0	0	0	0	1	0	0	0	0	1
Correct	6	5	1	16	1	1	4	0	0	0	0	28
Hustle	7	4	0	4	0	0	0	1	0	0	0	9
Scold	8	1	0	1	0	0	0	0	0	0	0	2
Manage	9	5	0	8	0	0	0	0	0	167	0	180
Other	10	1	0	2	0	0	0	0	0	2	5	10
TOTAL		555	87	493	1	1	28	9	2	180	10	1366
%		40.6	6.4	36.1	0.1	0.1	2.0	0.7	0.1	13.2	0.7	100

APPENDIX F

A SUGGESTED CODING SHEET FORMAT

As a considerable amount of time can be saved by using a computer to analyze CBOS generated data it is suggested that consideration be given to the coding sheet format in order to facilitate IBM key punching.

In this study the first 18 columns of the IBM card were reserved for identification. This I.D. included the

card number (3 cols,),
coding date (6 cols.),
coach (1 col.),
sample number (1 col.),
filming date (6 cols.),
sport (1 col.).

As key punchers are trained to punch data in horizontal rows, the coding sheet should conform accordingly. Thus, the coding sheet should contain twelve 5 second interval cells arranged in rows corresponding to the number of minutes of observation to be included on one sheet. In this study, thirty minutes of coding was included on one coding sheet, which resulted in thirty rows of twelve spaces each.

It is also necessary to include space on the coding sheet for other information such as the tape counter 000 start and stop, explanatory comments, and anything else that is required. Smith (1978) provides guidelines for this additional information.

In this study, the keypuncher was instructed to punch one row of the coding sheet onto one IBM card; that is, the first 18 columns were occupied by the ID, the

remaining 62 columns contained up to 31 (2 cols per digit) separately coded numbers in the sequence in which the behaviors were recorded. In most cases 31 entries adequately covered the numbers coded in any one row. However, on several occasions (where at times up to four different behaviors were coded in a five second interval) it was necessary to carry over to the next card.

APPENDIX G

CALCULATION OF SCOTT'S COEFFICIENT

1. Begin by tallying the behaviors coded into categories for a limited sample of the data (in this study five minute samples were used).
2. Place the totals for each of the ten categories in column 1 for observer 1 and column 2 for observer 2.
3. In column 3 place the percentages for observer 1; in column 4 percentages for observer 2.
4. Column 5 contains for each category the absolute (ignore signs) percent difference between the two observers.
5. The total of column 5 represents percent disagreement which is subtracted from 100 to give percent agreement (Po).
6. Take the average percent (both observers combined) figure for each category, square it and divide by 100, then enter these in column 6.
7. The sum of column 6 represents percent agreement expected by chance (Pe).
8. Scott's coefficient is calculated by:

$$\pi = \frac{Po-Pe}{100 - Pe}$$

APPENDIX G cont'd

EXAMPLE:

	1	2	% 1	% 2	% dis.	$\Sigma\%^2/100$
	32	41	31.4	36.3	4.9	11.4
	16	16	15.7	14.2	1.1	2.2
	43	47	42.2	41.6	0.6	17.5
	0	0	0.0	0.0	0.0	0.0
	0	0	0.0	0.0	0.0	0.0
	6	5	5.9	4.4	1.5	0.3
	2	0	2.0	0.0	2.0	0.0
	0	0	0.0	0.0	0.0	0.0
	1	2	1.0	1.8	0.8	0.0
	2	2	2.0	1.8	0.2	0.0
TOTAL	102	113	100	100	11.0	31.5

$$\pi = \frac{89 - 31.5}{100 - 31.5}$$

$$= \frac{57.5}{68.5}$$

$$\pi = 0.84$$